

**UNITED STATES DISTRICT COURT
FOR THE NORTHERN DISTRICT OF ILLINOIS**

KLEEN PRODUCTS LLC, individually and
on behalf of all others similarly situated,

Plaintiffs,

vs.

PACKAGING CORPORATION OF
AMERICA, et al.,

Defendants.

Case No. 1:10-cv-05711

Judge Milton I. Shadur

Magistrate Judge Nan R. Nolan

**DECLARATION OF MICHAEL J. FREED IN SUPPORT OF PLAINTIFFS'
MEMORANDUM OF LAW FOR EVIDENTIARY HEARING**

I, Michael J. Freed, declare as follows:

1. I am a partner with Freed Kanner London & Millen LLC and I am interim co-lead counsel for Plaintiffs in the above-captioned matter. I submit this declaration based on personal knowledge.
2. Attached hereto as **Exhibit A** is a true and correct copy of Plaintiffs' First Request for Production of Documents Directed to All Defendants, dated May 3, 2011.
3. Attached hereto as **Exhibit B** is a true and correct copy of Plaintiffs' Notice of Rule 30(b)(6) Depositions, dated November 2, 2011.
4. Attached hereto as **Exhibit C** is a true and correct copy of the curriculum vitae of Dr. David D. Lewis
5. Attached hereto as **Exhibit D** is a true and correct copy of the resume of Timothy D. Hanners.

6. Attached hereto as **Exhibit E** is a true and correct copy of an excerpt of a January 4, 2012 hearing transcript in *Moore v. Publicis Groupe*, No. 11-CV-1279 (S.D.N.Y.) (Peck, J.).

7. Attached hereto as **Exhibit F** is a true and correct copy of an article by Howard Turtle: “Natural Language vs. Boolean Query Evaluation,” *Proceedings of the 17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 212-220 (ACM, New York, 1994).

8. Attached hereto as **Exhibit G** is a true and correct copy of an article by the Hon. Andrew Peck: “*Search, Forward*,” Law Technology News (October 2011).

I declare under penalty of perjury under the laws of the United States that the foregoing is true and correct.

Dated: February 6, 2012

/s/ Michael J. Freed

Michael J. Freed

Exhibit

A

UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF ILLINOIS
EASTERN DIVISION

KLEEN PRODUCTS LLC, *et al.*, individually and on behalf of all others similarly situated,

Plaintiffs,

v.

PACKAGING CORPORATION OF AMERICA, *et al.*,

Defendants.

No. 1:10-cv-05711

PLAINTIFFS' FIRST REQUEST FOR PRODUCTION OF DOCUMENTS DIRECTED TO ALL DEFENDANTS

Pursuant to Rules 26 and 34 of the Federal Rules of Civil Procedure, Plaintiffs hereby request that each Defendant produce the documents and things herein requested within 30 days of the date of service at the offices of Freed Kanner London & Millen LLC, 2201 Waukegan Road, Suite 130, Bannockburn, IL 60015 A, or such other time and place as may be agreed upon by counsel. Pursuant to F.R.C.P. 34(b), unless otherwise specified, all documents must be organized and labeled to correspond to the categories in the associated document request. Corrections and supplemental answers and production of documents are required as provided for in the Federal Rules of Civil Procedure.

DEFINITIONS

The following definitions shall apply to these document requests:

1. "And" and "or" are to be read interchangeably so as to give the broadest possible meaning to a particular request in which either or both is used.

2. “Capacity reduction” refers to any permanent or near-permanent mill, plant or facility closure, curtailment or reduction in output from your maximum production capacity, but not including “Downtime” (defined *below*).

3. “Communication” means, without limitation, the imparting or exchange of information, thoughts or opinions by any means including orally, in writing, by signs, signals or code, including oral, written or electronic communications, such as face-to-face meetings, electronic communications, emails, facsimiles, telephone communications, correspondence or other exchange of written or recorded information. The phrase “communication between” is defined to include instances where one party addresses the other party but the other party does not necessarily respond.

4. “Competitive Conditions” means actual or forecasted, projected or estimated conditions or trends relating to pricing, production (including output, capacity, capacity utilization, operating rates, facility closures, rationalization, downtime and/or production scheduling), supply, inventories, demand, sales, profitability or margins, market share (including mergers, acquisitions, joint ventures, divestitures, sales or transfers of assets, spin-offs or any other form of change of ownership or control or consolidation), competitors, fixed or variable costs, imports or exports in the market for Containerboard Products.

5. “Complaint” means the Consolidated and Amended Complaint for Violation of the Sherman Act, filed November 8, 2010.

6. “Contact information,” when used with reference to a natural person, means to state their:

- a. present or last known home address;
- b. present or last known business address;

- c. all current, former and last known home, mobile and business telephone numbers;
- d. all current, former and last known personal and business email user identifications and email address information (used for any business purposes);
- e. all current, former and last known home and business facsimile numbers;
- f. all current, former or last known home and business text, SMS, MMS or similar addresses; and
- g. all current, former or last known positions and business affiliations.

7. “Containerboard Products” means linerboard, corrugated medium, corrugated or containerboard sheets and products constructed from corrugated sheets including boxes and other containers and, where context appropriate, also refers to the Containerboard Products-related business, industry or markets.

8. “Corporate Communications” means without limitation communications by the responding parties’ corporate officers or employees responsible in whole or in part for its external communications or messaging (*i.e.*, to media, financial community, customer base, government, industry bodies and institutes and general public), including, but not limited to, public relations, investor relations, government relations or similar functions or activities.

9. “Corporate Functions” shall be construed broadly in the context of the Complaint and includes, but is not limited to, the following executive, managerial or informational policy, study or analysis roles or responsibilities relating to any of the following functions or subjects and Containerboard Products:

- a. production (including output, capacity, capacity utilization, operating rates, facilities closures, rationalization, downtime and/or production scheduling);

- b. supply and inventories;
- c. pricing;
- d. sales, marketing and distribution;
- e. intra-company transfers or transactions (including transfer pricing);
- f. swaps, exchanges and other inter-Defendant or manufacturer transactions;
- g. business strategy, planning and analysis;
- h. monitoring, forecasting or reporting of market or Competitive Conditions;
- i. mergers, acquisitions, joint ventures, divestitures, sales or transfers of assets, spin-offs or any other form of change of ownership or control;
- j. finance and accounting, including financial performance and objectives;
- k. antitrust training or compliance;
- l. Corporate Communications; and
- m. purchasing or procurement of inputs, raw materials or intermediates (*i.e.*, pulp, OCC, linerboard, corrugated medium, dyes and inks, energy, supply chain, etc.) used in production.

10. “Defendant” means any company, organization, entity or person presently or subsequently named as a Defendant in this litigation including its predecessors, wholly-owned or controlled subsidiaries or affiliates, successors, parents, other subsidiaries, departments, divisions, joint ventures, other affiliates and any organization or entity managed or controlled by a named Defendant, including those merged with or acquired, together with all present and former directors, officers, employees, agents, attorneys, representatives or any persons acting or purporting to act on behalf of a Defendant.

11. "Document" shall be broadly interpreted as used in Rule 34(a) of the Federal Rules of Civil Procedure and means the original and all non-identical copies of all written, printed, recorded or electronically-stored information ("ESI"), including, but not limited to, letters, correspondence, memoranda, drafts, legal pleadings, calendars, diaries, day planners, journals, travel records, lists, outlines, summaries, notes, reports, compilations, notebooks, work papers, graphs, charts, spreadsheets, databases, blueprints, books, pamphlets, brochures, circulars, manuals, instructions, ledgers, drawings, sketches, telephone records, voicemail, telegrams, facsimiles, photographs, videotapes, audiotapes, digital audio or video recordings, other film and sound reproductions, presentations, slideshows, email, email attachments, text messages, instant messages, internal or external web sites, social networking data and messages, compact discs, computer files and disks, sales, advertising and promotional literature, agreements, stored recordings, minutes or other records of meetings, all written or graphic records or representations of any kind, and all electronic data, records or representations of any kind, including all file folders, directories, and metadata.

12. "Downtime" means any non-permanent mill curtailment, reduction in output or reduction in production capacity, including but not limited to, plant or facility closures, curtailment or production downtime imposed for maintenance (whether scheduled or unscheduled), curtailment or production downtime imposed to balance supply with demand, or reduction in output for any other reason.

13. "Electronically stored information" ("ESI") includes, without limitation, the following:

- a. activity listings of electronic mail receipts and/or transmittals;

b. output resulting from the use of any software program, including, without limitation, word processing documents, spreadsheets, database files, charts, graphs and outlines, electronic mail, Blackberry Messenger (or similar program including mobile instant messengers, web browser-based messengers or social network messengers) or bulletin board programs, operating and backup systems, source code, PRF files, PRY files, batch files, ASCII files, and all miscellaneous media on which they reside and regardless of whether said electronic data exists in an active file, a backup file or system, a deleted file or system, or file fragment; and

c. any and all items stored on computer memories, hard disks, floppy disks, CD-ROM, magnetic tape, microfiche, cloud data center(s) or in any other vehicle for digital data storage or transmittal, such as, but not limited to, desktop computers, servers and other network computers, backup tapes or systems, laptop computers, home or personal computers used for business purposes, a personal digital assistant, *e.g.*, Palm Pilot, R.I.M., Blackberry, iPhone or similar device (generically known as “smart phones” or “PDAs”), external storage devices (such as “keychain” drives) and file folder tabs, or containers and labels appended to, or relating to, any physical storage device associated with each original or copy of all documents requested herein.

14. “Employee” means, without limitation, any Person that you employ or employed during the Relevant Time Period.

15. “Identity” or “identify” means:

a. when used with reference to a natural person, to state his or her full name, and if known, his or her present home address, present business address, present home

and business telephone numbers, present or last known position and business affiliation and contact information;

b. when used in reference to any entity, such as a partnership, joint venture, trust or corporation, to state the full legal name of such entity, each name under which such entity does business, the entity's street address, the entity's telephone number, the identity of the chief operating officer, manager, trustee or other principal representative and the identity of those persons employed by or otherwise acting for such entity who are known or are believed to possess the knowledge or information responsive to the document request and for which the entity was identified; and

c. when used with reference to documents, to state specifically:

- i. the type of documents involved (e.g., letter, inter-office memorandum, etc.), together with information sufficient to enable a person to locate the document, such as its date, the name of any addressee, the name of any signer, the title or heading of the documents and its approximate number of pages; and
- ii. the identity of the person last known to have possession of the document, together with the present or last known location of the document. If any document was, but is no longer, in your possession, custody or control, identify the document, state what disposition was made of it and the reason for such disposition. In lieu of identifying documents, copies may be provided.

16. “Including” is used to emphasize certain types of documents requested and should not be construed as limiting the request in any way. Including therefore means “including, but not limited to,” or “including without limitation.”

17. “Industry Meeting” means, without limitation, any meeting, directly or indirectly relating to Containerboard Products or related products or industries, Trade Association (or other industry or similar organizations) or other entities or persons, whether formal, informal or *ad hoc*, involving two or more manufacturers, distributors or re-sellers of Containerboard Products or related businesses or persons and includes any formal or informal committee or other sub-unit.

18. “Management” shall be construed broadly and means without limitation, Employees or other Persons with executive, management, supervisory or policy-setting responsibilities, both individually and collectively, and, depending upon context, includes their administrative assistants, secretaries or other support staff.

19. “Manufacture,” “Manufacturing,” “Produce” or “Production” in each of their forms mean, without limitation, to make goods or products by manual labor or machinery or both, including to assemble, fabricate, construct, convert or process.

20. “Meeting” means, without limitation, any assembly, convocation, encounter, communication or contemporaneous presence (whether in person or via any electronic, computer-assisted, digital, analog, video or telephonic method of communication) of two or more persons for any purpose, whether planned or arranged, scheduled or not.

21. “Named Plaintiffs” means the following: Kleen Products LLC; R.P.R. Enterprises, Inc.; Mighty Pac, Inc.; Ferraro Foods of North Carolina, LLC; Distributors Packaging Group, LLC; RHE Hatco, Inc.; and Chandler Packaging, Inc.

22. “Person” means, without limitation, any individual, corporation, partnership or any variation thereof (e.g., limited partnership, limited liability partnership), limited liability company, proprietorship, joint venture, association, group or other form of legal entity or business existing under the laws of the United States, any state or any foreign country.

23. “Planning Staff” shall be construed broadly and means without limitation, Employees or other Persons with strategic planning, informational or analysis responsibilities both individually and collectively, and, depending upon context, includes their administrative assistants, secretaries or other support staff.

24. “Policy” or “procedure” means any rule, practice or course of conduct, guidelines or business methods or traditions whether formal or informal, written or unwritten, recorded or unrecorded, which was recognized or followed, explicitly or implicitly, by you.

25. “Relating to,” “referring to,” “regarding,” “with respect to” or “concerning” mean without limitation the following concepts: concerning, constituting, discussing, describing, reflecting, dealing with, pertaining to, analyzing, evaluating, estimating, constituting, studying, surveying, projecting, assessing, recording, summarizing, criticizing, reporting, commenting, or otherwise involving, in whole or in part, directly or indirectly. Documents are considered relating to the subject matter whether they are viewed alone or in combination with other documents.

26. “Relevant Time Period” means January 1, 2002 through and including December 31, 2010, unless otherwise noted.¹

27. “Sales Personnel” shall be construed broadly and means, without limitation, any Employees or other Persons that had responsibility, directly or indirectly, relating to sales to customers of Containerboard Products sold in the U.S. (including any role or responsibility for

¹ See Instruction No. 1, *infra*.

selling, determining or establishing pricing, effectuating, or establishing terms and conditions of sales or purchases).

28. “Studies” or “analyses” include all reports, memoranda, statistical compilations, reviews, audits and other types of written, printed, or electronic submissions of information.

29. “Trade Association” means any group or association, formal or informal, relating to the market for Containerboard Products, including the Association of Independent Corrugated Converters (“AICC”), the Fibre Box Association (“FBA”), The American Forest & Paper Association (“AF&PA”), The International Corrugated Case Association (“ICCA”), Corrugated Packaging Alliance (“CPA”), the Corrugated Packaging Alliance Action Team (“CPAAT”), the American Forest & Paper Association (“AFPA”), the International Corrugated Case Association (“ICCA”), the Paper Industry Management Association (“PIMA”), the Association of Independent Corrugated Converters (“AICC”) and RISI (originally founded and known as Resource Information Systems, Inc.) and includes any formal or informal committee or other sub-unit.

30. “United States” or “U.S.” means each of the 50 states, the District of Columbia, and the commonwealths, territories, and possessions of the United States.

31. “You,” “your” or “your company” mean the responding Defendant, including its predecessors, wholly-owned or controlled subsidiaries or affiliates, successors, parents, other subsidiaries, departments, divisions, joint ventures, other affiliates and any organization or entity that the responding Defendant manages or controls, including those merged with or acquired, together with all present and former directors, officers, employees, agents, attorneys, representatives or any persons acting or purporting to act on their behalf.

INSTRUCTIONS

1. Each request seeks all responsive documents created or generated during the designated time period, as well as all responsive documents created or generated outside the designated time period that contain information concerning the designated time period. Changes that occurred within a designated time period shall be included in the documents produced.

2. Pursuant to Rule 26(e) of the Federal Rules of Civil Procedure, these document requests are continuing in nature so that if you subsequently discover or obtain possession, custody or control of any document covered by these requests, you shall promptly make any such document available to Plaintiffs.

3. In producing documents and other materials, you are to furnish all documents in your possession, custody or control regardless of whether such documents or materials are physically or otherwise possessed, in the custody of or controlled directly or indirectly by you or your employees or former employees, agents or former agents, parents, subsidiaries, affiliates, investigators or by your attorneys or their employees, agents or investigators. (Archived documents, ESI or other information maintained by third-party vendors, cloud data centers, or off-site shall be considered within your possession, custody or control.)

4. Pursuant to Rule 34(b) of the Federal Rules of Civil Procedure, unless otherwise specified, all documents must be organized and labeled to correspond to categories in the associated document request. For purposes of this instruction, the labels in Exhibit "A" shall be used as a "coding field" entry in the parties' document production.

5. All documents shall be produced in the file folder, envelope or other container in which the documents are kept or maintained. If for any reason the container cannot be produced,

you should produce copies of all labels or other identifying marks that may be present on the container.

6. Documents attached to one another should not be separated. If any portion of any document is responsive to any portion of the document requests below, then the entire document must be produced.

7. Documents shall be produced in such fashion as to identify the department, branch or office in whose possession they were located and, where applicable, the natural person in whose possession they were found (*i.e.*, the document custodian) and the business address of each document custodian.

8. All documents produced should be numbered sequentially, with a unique production number on each page (*i.e.*, a “bates” number), and with a prefix identifying the party producing the document. An electronic document produced in its native format shall include the document’s production number and an MD5 or similar hash number as associated metadata.

9. To distinguish between paper and electronic records, all paper records shall include the letter “P” in the applicable Bates prefix.

10. If a document once existed and subsequently has been lost, destroyed or is otherwise missing, you should provide sufficient information to identify the document and state, in writing, the details, including whether the document:

- a. is lost or missing;
- b. has been destroyed and, if so, by whom at whose request;
- c. has been transferred or delivered, voluntarily or involuntarily, to another person or entity and at whose request; or
- d. has been otherwise disposed of.

11. In each instance in which a document once existed but now is lost or missing or has been destroyed or otherwise disposed of, explain the circumstances surrounding the disposition of the document, including, but not limited to:

- a. the identity of the person or entity who last possessed the document;
- b. the date or approximate date of the document's disposition; and
- c. the identity of all persons who have or had knowledge of the documents' contents.

12. If any document responsive to any of these requests is privileged, and the document or any portion of the document requested is withheld based on a claim of privilege pursuant to Rule 26(b)(5) of the Federal Rules of Civil Procedure, provide a statement of the claim of privilege and all facts relied upon in support of that claim, as required by Federal Rules of Civil Procedure, including the following information:

- a. the reason for withholding the document;
- b. the date of such communication;
- c. the medium of such communication;
- d. the general subject matter of such communication (such description shall not be considered a waiver of your claimed privilege);
- e. the identity of any document that was the subject of such communication and the present location of any such document;
- f. the identity of all the persons involved in such communication;
- g. the identity of any document which records, refers, or relates to such communication and present location of any such document; and

h. the number or numbers of these requests for production of documents to which such information is responsive.

13. Each document requested herein should be produced in its entirety and without deletion, redaction or excision, except as qualified by Instruction 11 above, regardless of whether you consider the entire document or only part of it to be relevant or responsive to these document requests. If you have redacted any portion of a document on the ground of privilege, stamp the word “REDACTED” beside the redacted information on each page of the document which you have redacted. Any redactions to documents produced should be identified in accordance with Instruction 11 above.

14. All documents produced in paper form should be numbered sequentially, with a unique number on each page, and with a prefix identifying the party producing the document.

15. All ESI shall be produced in such fashion as to identify the location, *i.e.*, the network file folder, hard drive, back-up tape, or other location, where the documents are stored and, where applicable, the natural person in whose possession they were found and the business address of each document’s custodian(s). All ESI must include a .dii file or similar industry standard load file indicating the beginning and ending Bates numbers of all documents.

16. ESI shall be produced, in its native format without deletion or alteration of metadata, or in a reasonably usable format agreed upon by the parties and shall include information or instructions sufficient to enable to the propounding party to access the ESI in such format.

17. If a producing party produces electronic documents without some or all of the metadata that was contained in the electronic documents, the producing party shall inform all other parties of this fact, in writing, at or before the time of production.

SPECIFIC DOCUMENT REQUESTS²

Organizational Structure

REQUEST NO. 1:

For the period January 1, 2000 through December 31, 2010, organizational charts or other documents sufficient to show your complete organizational structure directly or indirectly relating to Containerboard Products, including a description of the business functions or responsibilities of each of the following:

- a. predecessors, parents, subsidiaries, affiliates and joint ventures,
- b. divisions, departments, segment, units or sub-divisions, and
- c. board or management committees, subcommittees, or working groups and

each of their related Persons.

Document Preservation

REQUEST NO. 2:

For the period January 1, 2000 to the present, all documents relating to your policies or practices concerning the backup, archiving, retention, storage, destruction, disposal, or preservation of documents or ESI (including litigation hold policies and procedures).

Current and Former Personnel

REQUEST NO. 3:

Documents sufficient to identify:

- a. the names, titles, job descriptions, business affiliations³ and contact information of each of your current or former directors and officers (including each of their administrative assistants, secretaries or other support staff);

² Sub-headings and footnotes are included for convenience only and are qualified in their entirety by the relevant Definitions, Instructions and Specific Requests.

- b. the names, titles and job descriptions, business affiliations and contact information of each of your current or former Management with responsibilities relating to the Corporate Functions (including each of their administrative assistants, secretaries or other support staff);
- c. the names, titles and job descriptions business affiliations and contact information each of your current or former Planning Staff with responsibilities relating to the Corporate Functions (including each of their administrative assistants, secretaries or other support staff);
- d. the names, titles and job descriptions, business affiliations and contact information each of your current or former Sales Personnel (including each of their administrative assistants, secretaries or other support staff);
- e. the names, titles and job descriptions business affiliations and contact information each of your current or former directors, officers, Employees or vendors responsible for maintaining any ESI or other electronic database(s), including archives, of email or other electronic documents relating to the Corporate Functions (including retention and/or destruction policies or implementation); and
- f. the names, titles and job descriptions business affiliations and contact information each of your current or former directors, officers, Employees or vendors responsible for maintaining any central files or archives of paper document relating to the Corporate Functions (including retention and/or destruction policies or implementation).

³ For purposes of this Request, “business affiliations” includes affiliations with any of you, your predecessors, parents, subsidiaries, affiliates, joint ventures, divisions, departments, segment, units or sub-divisions, board or management committees, subcommittees, or working groups. “Business affiliations” also includes known business relationships with any other Persons.

REQUEST NO. 4:

For each your current or former directors and officers, Management, Planning Staff and Sales Personnel identified in the foregoing Request No. 3 (excluding administrative assistants, secretaries or other support staff) all:

- a. electronic and manual diaries, calendars, appointment books, to do lists, Day Timers, day planners or appointment notes;
- b. trip and travel logs, records, expenses, and other supporting documents;
- c. expense or entertainment reports, including supporting documents;
- d. personal and company copies of telephone number logs, directories, notebooks, Rolodex cards and related memoranda;
- e. bills, statements, records, and supporting documents concerning local, long distance, or cellular telephone calls by such employees, including calls made using telephones not paid for by you (such as home telephones or personal cellular telephones) if such telephones were used for business purposes;
- f. documents relating to membership in any Trade Association or industry group;
- g. contact information including all email addresses used by such persons for any business purposes, even if only sparingly; and
- h. communications with any other Defendant or their Employees or agents.

REQUEST NO. 5:

All documents relating to the voluntary or involuntary termination, resignation, retirement, discipline, discharge, reassignment, transfer or suspension of any current or former officers, directors or Employees with responsibilities relating to the Corporate Functions as a

result or relating to any contacts or communications with directors, officers, Employees, or any other representatives of any other Defendant or producer of Containerboard Products.

REQUEST NO. 6:

All documents relating to the voluntary or involuntary termination, resignation, retirement, discipline, discharge, reassignment, transfer or suspension of any current or former officers, directors, Employees or other Person as a result of any alleged participation in, involvement in or knowledge of any agreement, understanding, plan, contract, combination or conspiracy, express or implied, direct or indirect, in any manner relating to the production, marketing, distribution, sale, or resale of Containerboard Products.

Trade Associations & Industry Meetings

REQUEST NO. 7:

All documents relating to any Trade Associations (including membership lists, meeting announcements, agendas, minutes, notes, attendance lists, expense reports, handouts, reports, statistical bulletins, or correspondence).

REQUEST NO. 8:

Documents sufficient to identify any of your current or former officers, directors, Employees or other Persons acting on your behalf who attended any Trade Association meeting including the dates, locations, subject matter and sponsors of any such meeting(s).

REQUEST NO. 9:

All documents relating to any Industry Meetings (including meeting announcements, agendas, minutes, notes, attendance lists, expense reports, handouts, reports, statistical bulletins, or correspondence).

REQUEST NO. 10:

Documents sufficient to identify any of your current or former officers, directors, Employees or other Persons acting on your behalf who attended any Industry Meetings including the dates, locations, subject matter and sponsors of any such meeting(s).

REQUEST NO. 11:

For any other Person not covered by Request Nos. 3 or 4 but who attended any Trade Association meetings or events or Industry Meeting(s) all:

- a. electronic and manual diaries, calendars, appointment books, to do lists, Day Timers, day planners or appointment notes;
- b. trip and travel logs, records, expenses, and other supporting documents;
- c. expense or entertainment reports, including supporting documents;
- d. personal and company copies of telephone number logs, directories, notebooks, Rolodex cards and related memoranda;
- e. bills, statements, records, and supporting documents concerning local, long distance, or cellular telephone calls by such employees, including calls made using telephones not paid for by you (such as home telephones or personal cellular telephones) if such telephones were used for business purposes;
- f. documents relating to membership in any Trade Association or industry group;
- g. contact information including all email addresses used by such persons for any business purposes, even if only sparingly , and
- h. communications with any other Defendant or their Employees or agents.

Products & Manufacturing Processes

REQUEST NO. 12:

For the period January 1, 2000 through December 31, 2010, documents sufficient to identify the countries, geographic regions or territories where you manufacture, sell, market, or distribute Containerboard Products including each class, type, grade, or category.

REQUEST NO. 13:

For the period January 1, 2000 through December 31, 2010, documents sufficient to identify each class, type, grade, or category of Containerboard Product that you marketed, sold or distributed in the United States, including any brand name, its function, its physical properties and characteristics (including flute designation, flutes per linear foot, flute thickness per inch, flute per linear meter, other measures of flute thickness and ATSM designation).⁴

REQUEST NO. 14:

For the period January 1, 2000 through December 31, 2010, documents sufficient to identify each class, type, grade, or category of Containerboard Product you manufactured in the United States, including any brand name, its function, its physical properties and characteristics including flute designation, flutes per linear foot, flute thickness per inch, flute per linear meter, other measures of flute thickness and ATSM designation.⁵

REQUEST NO. 15:

All documents that compare or contrast Containerboard Products sold by you in the U.S with Containerboard Products sold by any other Defendant or by any other manufacturer,

⁴ Containerboard Products that are marketed, sold or distributed in the U.S are included within this Request regardless of the country where they were actually manufactured.

⁵ Containerboard Products that are manufactured in the U.S. are included within this Request regardless of the county where they marketed, sold or distributed.

distributor, re-seller or broker (including product characteristics, specifications, quality, prices, availability, delivery schedules, etc.).

REQUEST NO. 16:

All documents relating to actual or potential substitutes for Containerboard Products (including the effect of such substitutes' prices upon Containerboard Products pricing, purchase terms or profits).

REQUEST NO. 17:

For the period January 1, 2000 through December 31, 2010, documents sufficient to show the production process of each class, type, grade or category of Containerboard Products that you manufactured, marketed, sold or distributed in the U.S. (including raw material inputs or intermediates and the types of sheet, corrugating or converting machines used).

Mills, Plants & Converting Facilities

REQUEST NO. 18:

For the period January 1, 2000 through December 31, 2010, documents sufficient to determine the identity of all mills, plants, converting facilities or other facilities where you produced Containerboard Products sold in the U.S., including the identity of each sheet, corrugating or converting machine at each particular facility.⁶

REQUEST NO. 19:

For the period January 1, 2000 through December 31, 2010, documents sufficient to determine the dates that production of Containerboard Products sold in the U.S. began at each mill, plant, converting facility or other facilities, including:

- a. production starting dates for each particular facility;

⁶ This Request includes all Containerboard Products that you sold in the U.S. regardless of the country of manufacturing origin.

- b. production starting dates for each particular sheet, corrugating or converting machine at that facility; and
- c. production starting dates for each brand, class, type, grade, or category of Containerboard Product manufactured at that facility.

REQUEST NO. 20:

For the period January 1, 2000 through December 31, 2010, for each mill, plant, converting facility or other facility where you produced Containerboard Products that you marketed, sold or distributed in the U.S, documents sufficient to:

- a. identify the subsidiaries, affiliates, joint ventures, divisions, departments, segments, units or sub-divisions or other Persons that owned, operated or controlled the particular facility;
- b. identify (by matching) each brand, class, type, grade, or category of Containerboard Product manufactured at that particular facility; and
- c. identify (by matching) each brand, class, type, grade, or category of Containerboard Product manufactured to each particular sheet, corrugating or converting machine at that particular facility.

Production Capacity⁷

REQUEST NO. 21:

For the period January 1, 2000 through December 31, 2010, for each mill, plant, converting facility or other facility where you produced Containerboard Products that you marketed, sold or distributed in the U.S, documents sufficient to determine the weekly, monthly,

⁷ Produce the requested documents on a facility-by-facility, product-by-product (including class, type, grade, or category), machine-by-machine and product-by-machine basis or that matches the facilities, products and machines as they are used in the production process.

quarterly and annual (i) rated production capacity (according to generally accepted rating methodologies, such as thousands of short tons) and (ii) actual production capacity (if different from rated capacity, including the method or metrics of measurement and the means used to calculate such capacity). The requested information must be provided for (i) each particular facility; (ii) each particular sheet, corrugating or converting machine at that facility; and (iii) producing each brand, class, type, grade, or category of Containerboard Product manufactured at that facility.

REQUEST NO. 22:

For the period January 1, 2000 through December 31, 2010, for each mill, plant, converting facility or other facility where you produced Containerboard Products that you marketed, sold or distributed in the U.S., all documents that relate to any changes or variance in weekly, monthly, quarterly and annual rated or actual production capacity including all reasons for such changes or variance. The requested information must be provided for (i) each particular facility; (ii) each particular sheet, corrugating or converting machine at that facility; and (iii) producing each brand, class, type, grade, or category of Containerboard Product manufactured at that facility.

Operating Rates, Capacity Changes, Downtime⁸

REQUEST NO. 23:

For the period January 1, 2000 through December 31, 2010, for each mill, plant, converting facility or other facility where you produced Containerboard Products that you marketed, sold or distributed in the U.S., documents sufficient to determine actual weekly,

⁸ Produce the requested documents on a facility-by-facility, product-by-product (including class, type, grade, or category), machine-by-machine and product-by-machine basis or in any way that matches the facilities, products and machines as they are used in the production process.

monthly, quarterly and annual (i) forecasted, projected or estimated operating or capacity utilization rates and (ii) actual operating or capacity utilization rates. The requested information must be provided for (i) each particular facility; (ii) each particular sheet, corrugating or converting machine at that facility; and (iii) producing each brand, class, type, grade, or category of Containerboard Product manufactured at that facility.

REQUEST NO. 24:

For the period January 1, 2000 through December 31, 2010, for each mill, plant, converting facility or other facility where you produced Containerboard Products that you marketed, sold or distributed in the U.S., all documents that relate to changes or variance in (i) the forecasted, projected or estimated weekly, monthly, quarterly and annual operating or capacity utilization rates and (ii) the actual weekly, monthly, quarterly and annual operating or capacity utilization rates, including the reasons for such changes or variances. The requested information must be provided for (i) each particular facility; (ii) each particular sheet, corrugating or converting machine at that facility; and (iii) producing each brand, class, type, grade, or category of Containerboard Product manufactured at that facility.

REQUEST NO. 25:

For the period January 1, 2000 through December 31, 2010, documents sufficient to determine all Capacity Reductions, capacity additions or other changes to capacity, by each mill, plant, converting facility or other facility where you produced Containerboard Products that you marketed, sold or distributed in the U.S, including both planned or actual changes. The requested information must be provided for (i) each particular facility; (ii) each particular sheet, corrugating or converting machine at that facility; and (iii) producing each brand, class, type, grade, or category of Containerboard Product manufactured at that facility.

REQUEST NO. 26:

For the period January 1, 2000 through December 31, 2010, all documents that relate to any reasons for any Capacity Reductions, capacity additions or changes to capacity by each mill, plant, converting facility or other facility where you produced Containerboard Products that you marketed, sold or distributed in the U.S, including both planned or actual changes. The requested information must be provided for (i) each particular facility; (ii) each particular sheet, corrugating or converting machine at that facility; and (iii) producing each brand, class, type, grade, or category of Containerboard Product manufactured at that facility.

REQUEST NO. 27:

For the period January 1, 2000 through December 31, 2010, documents sufficient to determine all Downtime of each mill, plant, converting facility or other facility where you produced Containerboard Products that you marketed, sold or distributed in the U.S. (including both scheduled and unscheduled downtime). The requested information must be provided for (i) each particular facility; (ii) each particular sheet, corrugating or converting machine at that facility; and (iii) producing each brand, class, type, grade, or category of Containerboard Product manufactured at that facility.

REQUEST NO. 28:

For the period January 1, 2000 through December 31, 2010, all documents that relate to any reasons for all Downtime incurred at each mill, plant, converting facility or other facility where you produced Containerboard Products that you marketed, sold or distributed in the U.S (including both scheduled and unscheduled Downtime). The requested information must be provided for (a) each particular facility; (b) each particular sheet, corrugating or converting

machine at that facility; and (c) producing each brand, class, type, grade, or category of Containerboard Product manufactured at that facility.

Pricing

REQUEST NO. 29:

All documents utilized or generally referred to by your Management or Sales Personnel for quoting, changing, or setting the prices or the terms and conditions for sale of Containerboard Products (including documents relating to pricing guidelines, pricing methods, pricing formulas, procedures or authorization procedures for Sales Personnel to quote a price, price changes, price lists, pricing policies, pricing guidelines, pricing methods, violations of pricing policies and documents providing guidance to Sales Personnel about implementation of price changes).

REQUEST NO. 30:

All documents relating to or communicating price changes of Containerboard Products produced or sold in the U.S., (including price announcements, explanations of the reasons for price changes).

REQUEST NO. 31:

All documents analyzing the prices of Containerboard Products produced or sold in the U.S.

REQUEST NO. 32:

All documents relating to the forecasted, projected, estimated or actual effects of price increases for Containerboard Products (including any documents relating to any evaluation of the success or failure of any price increase).

REQUEST NO. 33:

For the period January 1, 2000 through December 31, 2010, all price lists or similar documents relating to Containerboard Products produced and/or sold in the U.S.

REQUEST NO. 34:

All non-transaction specific documents, including communications, relating to the price of Containerboard Products offered by you, by any other Defendant, or by any other manufacturer, distributor, re-seller or broker of Containerboard Products (including price announcements, price changes, explanations of the reasons for price changes, price lists, pricing policies, pricing guidelines, pricing methods, pricing formulas, procedures or authorization procedures for Sales Personnel to quote a price, analysis of market prices, and monitoring of competitor pricing).

REQUEST NO. 35:

For the period January 1, 2000 to the present, documents sufficient to show weekly, monthly, quarterly, and annual prices of Containerboard Products sold in the U.S. including prices offered by you, by any other Defendant, or by any other manufacturer, distributor, re-seller or broker of Containerboard Products identified on a (i) product line or product category basis; and (ii) product line-by-facility or product category-by-facility basis.

REQUEST NO. 36:

All documents relating to the elasticity of demand or price sensitivity of Containerboard Products.

REQUEST NO. 37:

All documents relating to any index or indices prices.

Cost Accounting

REQUEST NO. 38:

Documents sufficient to identify the names, titles, job descriptions, business affiliations and contact information of each of your current or former Employees with executive or management responsibility for cost accounting (production, sales, and distribution costs, etc.) for Containerboard Products manufactured or sold in the U.S. including on a (i) consolidated basis; (ii) for each business unit or segment (iii) product line or product category basis; and (iv) facility-by-facility basis.

REQUEST NO. 39:

Documents identifying all software, programs, databases and systems you used in connection with cost accounting for Containerboard Products manufactured or sold in the U.S., including their outputs of such software, programs, databases and systems.

REQUEST NO. 40:

Documents sufficient to identify all types of fixed and variable production, sales or marketing costs, including raw materials, that you track or record for Containerboard Products manufactured or sold in the U.S.

REQUEST NO. 41:

All documents identifying and relating to your cost accounting policies and procedures for Containerboard Products manufactured or sold in the U.S.

REQUEST NO. 42:

All documents relating to accounting policies and procedures for intra-company transfer pricing and intra-defendant transactions for Containerboard Products manufactured or sold in the U.S.

REQUEST NO. 43:

All weekly, monthly, quarterly and annual cost or cost accounting reports (including, production, sale and distribution) regularly prepared or received by your current or former directors, officers, Management, Planning Staff on a (i) consolidated basis; (ii) for each business unit or segment (iii) product line or product category basis; and (iv) facility-by-facility basis, including all documents or communications related to the foregoing.

REQUEST NO. 44:

All budgets, projections, estimates, or related studies or reports prepared or received by your current or former directors, officers, Management, or Planning Staff on a (i) consolidated basis; (ii) for each business unit or segment (iii) product line or product category basis; and (iv) facility-by-facility basis.

Competitive Conditions and Reports

REQUEST NO. 45:

All documents relating to Competitive Conditions in the market for Containerboard Products, including reports, studies or analyses relating to forecasted, projected, estimated, planned or actual: (i) market shares, (ii) consolidation, (iii) production capacity, capacity reduction, capacity utilization or operating rates (iv) fixed or variable costs, (v) pricing, (vi) inventories (vii) entry or exist conditions (viii) inventories, (ix) supplies/supply trends, (x) demand/demand trends, or (xi) substitute products.

REQUEST NO. 46:

All documents relating or referring to price discipline, price leadership, industry leadership, industry discipline, production discipline, supply discipline or a disciplined approach or practices.

REQUEST NO. 47:

For the period January 1, 2000 through December 31, 2010 all documents relating to monitoring, analyzing, or studying any competitor's: (i) market shares, (ii) pricing, (iii) production capacity, capacity reduction, capacity utilization or operating rates, (iv) downtime, (v) supplies and inventories, or (vi) customers.

REQUEST NO. 48:

All documents concerning Containerboard Products sold in the U.S. relating to policies, studies, reports, presentations received or prepared by your current or former directors, officers, Management or Planning Staff relating to the forecasted, projected, estimated, planned or actual:

- a. effects of production capacity, capacity reduction, capacity utilization, operating rates or downtime (including changes) on prices, supplies or inventories;
- b. effects of supplies or inventories (including changes) on prices; and
- c. effects of price changes on demand.

REQUEST NO. 49:

All documents relating to studies, reports and presentations received or prepared by your current or former directors, officers, Management or Planning Staff relating to Competitive Conditions in the market for Containerboard Products (including management reports and presentations, financial reports and presentations, business planning and strategy reports and presentations, planning analyses, budgets, forecasts, sales projections consulting and advisory reports and presentations).

REQUEST NO. 50:

All business or strategic plans relating to Containerboard Products, for the period January 1, 2000 to December 31, 2010.

REQUEST NO. 51:

All documents relating to forecasted, projected, estimated, planned, or actual demand for Containerboard Products.

REQUEST NO. 52:

All documents relating to demand changes or demand fluctuations for Containerboard Products sold in the U.S. (including the underlying causes of demand fluctuations and the effect of demand fluctuations upon Containerboard Products pricing, purchase terms or profits).

REQUEST NO. 53:

All documents relating to any memoranda, studies, reports, analyses, or presentations concerning Containerboard Products, prepared by any consulting firms (including McKinsey & Company, Deloitte Consulting LLP or Accenture).

REQUEST NO. 54:

All documents relating to any memoranda, studies, reports, analyses, or presentations concerning Containerboard Products, prepared by any investment banking, commercial banking or other financial services firms (including Deutsche Bank, Citigroup, Lazard Ltd., UBS or Morgan Stanley).

REQUEST NO. 55:

All documents relating to any trade publications, market analyst studies, reports, analyses, presentations, newsletters, publications, market updates or other similar documents concerning Containerboard Products (including documents authored by Deutsche Bank, Buckingham Research Group Inc., the Gerson Lehman Group, RISI, Official Board Markets or any other market analyst or firm).

REQUEST NO. 56:

All documents relating to memoranda, studies, reports or presentations relating to Containerboard Products that were delivered to investors, annual meeting attendees, earnings call participants, roadshow attendees, prospective investors, lenders, bondholders, other creditors or prospective creditors, investment banks, investment funds, commercial banks, market analysts, or credit rating agencies.

REQUEST NO. 57:

All documents relating to memoranda, studies, reports or presentations (including related documents such as speeches or speaker's notes) received or delivered by current or former current or former directors, officers, Management, Planning Staff or Sales Personnel who attended any Trade Association meeting, Industry Meeting or similar event attended by any other Defendant (including those sponsored by any Trade Association, consulting firm, investment or commercial bank, trade publication or market information provider or analyst).

REQUEST NO. 58:

All documents relating to studies, reports, presentations, newsletters or data reports concerning Containerboard Products that were submitted to, received from, or published by any Trade Association or other formal or informal industry group.

REQUEST NO. 59:

All documents relating to any meeting, communications (including corporate communications) with any financial analyst, investment or commercial bank, trade publication or market information provider or analyst concerning Competitive Conditions in the market for Containerboard Products.

REQUEST NO. 60:

All documents relating to any contemplated, proposed, or actual bids for Containerboard Products business, including documents concerning whether or not to bid.

Competitor Contacts

REQUEST NO. 61:

For the period January 1, 2000 through December 31, 2010, all documents that refer to, relate to, or contain information concerning, in whole or in part, any consummated, proposed, direct or indirect meeting, communication, written or unwritten agreement or understanding involving any of your current or former officers, directors or Employees or other Persons acting on your behalf, on the one hand, and any current or former officers, directors, Employees or other Persons acting on behalf of any other Defendant or other Person involved in production, marketing, distribution, sale, or resale of Containerboard Products on the other hand, relating in any manner to Containerboard Products:

- a. production (including output, capacity, capacity reduction, capacity utilization, operating rates, facilities closures, rationalization, downtime and/or production scheduling);
- b. supply and inventories;
- c. pricing (including prices, pricing and terms, discounts, conditions of sale, price increases, maintaining prices and price increases and strategy for implementing price increases);
- d. sales and marketing;
- e. customers;
- f. imports or exports;

- g. swaps, exchanges, intra-company transfers and inter-Defendant transactions;
- h. business strategy, planning and analysis;
- i. market or Competitive Conditions;
- j. mergers, acquisitions, joint ventures, divestitures, sales or transfers of assets, spin-offs or any other form of change of ownership or control;
- k. finance and accounting, including financial performance and objectives;
- l. antitrust training or compliance;
- m. Corporate Communications;
- n. purchasing or procurement or supply conditions of inputs, raw materials or intermediates (*i.e.*, pulp, OCC, linerboard, corrugated medium, dyes and inks, energy, supply chain, etc.) used in production;
- o. market share levels or changes in market share; and
- p. profit margins or profitability.

Industry Consolidation

REQUEST NO. 62:

For the period January 1, 2000 through December 31, 2010, all documents relating to any proposed or consummated sale, acquisition, merger, joint venture, divestiture, transfer of assets, spin-off, or any other form of change of ownership or control or business combination concerning any Containerboard Products related business, production facilities, product lines or other assets (including both businesses or assets owned or controlled by you and those owned or controlled by other Persons and any minutes, notes, reports, studies, analysis, presentations,

expert or market participant comments, analysis or other information and all such documents submitted to or received from any governmental agency).

Announcements, Press Releases & Corporate Communications

REQUEST NO. 63:

All documents relating to public announcements, press releases, or Corporate Communications concerning Containerboard Products that were made by you, or any other Defendant or any other manufacturer, distributor, re-seller or broker of Containerboard Products:

- a. production (including output, capacity, capacity reduction, capacity utilization, operating rates, facilities closures, rationalization, downtime and/or production scheduling);
- b. supply and inventories;
- c. pricing (including prices, pricing and terms, discounts, conditions of sale, price changes, maintaining prices and strategy for implementing price changes);
- d. sales and marketing;
- e. customers;
- f. imports or exports;
- g. swaps, exchanges, intra-company transfers and inter-Defendant transactions;
- h. business strategy, planning and analysis;
- i. Competitive Conditions;
- j. mergers, acquisitions, joint ventures, divestitures, sales or transfers of assets, spin-offs or any other form of change of ownership or control;
- k. finance and accounting (including financial performance and objectives);

- l. antitrust training or compliance;
- m. purchasing and procurement and supply condition of inputs, raw materials or intermediates (*i.e.*, pulp, OCC, linerboard, corrugated medium, dyes and inks, energy, supply chain, etc.) used in production;
- n. market share levels or changes in market share; and
- o. profit margins or profitability.

Inquiries, Investigation, Litigation & Compliance Policies, Procedures, Practices and Guidelines

REQUEST NO. 64:

For the period January 1, 1996 through December 31, 2010, all documents that directly or indirectly relate, in whole or in part, alone or when viewed with other documents any combination, conspiracy, agreement or understanding, explicit, implied or tacit, to fix prices, reduce output, production capacity supplies or inventories or allocate customers, product lines or territories for Containerboard Products sold in the U.S.

REQUEST NO. 65:

For the period January 1, 1996 through December 31, 2010, all documents relating to any possible or alleged violations of federal, state, or international antitrust laws, competition laws, anti-monopoly laws, anti-cartel laws, unfair competition laws, anti-bribery laws and any similar laws, rules or regulations by you, any other Defendant or any other Containerboard Products-related business or person (including any current or former officers, directors, Employees or other persons) including:

- a. any inquiry, investigation, administrative, criminal or civil litigation or similar actions (regardless of characterizations such as non-adjudicative, cooperating, preliminary, informal, formal, incomplete, open, closed or final) by any government

entity, legislative body or representative or instrumentality thereof (*i.e.*, the United States Department of Justice, the Federal Bureau of Investigation, the Federal Trade Commission, the International Trade Commission, any state law enforcement agency, any grand jury, any committee of the United States Congress, any agency of a state or other political subdivision, and any domestic or international government body) (including all documents received from or provided to such governmental entities or used during or regarding discussions or negotiations with them);

- b. any internal inquiry or investigation (regardless of characterizations such as non-adjudicative, cooperating, preliminary, informal, formal, incomplete, open, closed or final) and further including any violation of internal antitrust or competition law policies or practices; and
- c. any private actions or civil litigation (including arbitrations, mediations or negotiations and further including any violation of internal antitrust or competition law policies or practices) in which you have been a party, witness, participant, or subject.

REQUEST NO. 66:

For the period January 1, 1996 through December 31, 2010, all documents relating to actions to conceal or avoid detection of any potential violations of federal, state, or international antitrust laws, competition laws, anti-monopoly laws, anti-cartel laws, unfair competition laws, anti-bribery laws and any similar laws, rules or regulations (including meeting in locations such as coffee shops, airport hotels, and going for a walk outside, using non-traceable prepaid calling cards, non-contract or disposable cell phones, placing calls from public phones, and destroying, secreting, altering or forging documents).

REQUEST NO. 67:

Without time limitation, documents:

- a. sufficient to show the terms and conditions of any preliminary or permanent injunction, consent decree, plea agreement, sentencing agreement, settlement agreement or stipulation of settlement in a class action or private action, judgment or similar orders or agreements relating to alleged violations of federal, state, or international antitrust laws, competition laws, anti-monopoly laws, anti-cartel laws, unfair competition laws, anti-bribery laws and any similar laws, rules or regulations entered against you or any other Defendant, further including in the following actions:
 - i. *United States of America v. National Container Association, et al.;*
 - ii. *United States v. Container Corporation of America, et al.;*
 - iii. *In re Folding Carton Antitrust Litigation;*
 - iv. United States v. 14 companies and 26 individuals, including International Paper, Weyerhaeuser and Stone Container Corporation in United States District Court for the Southern District of Texas alleging a conspiracy east of the Rocky Mountains to fix prices of corrugated containers and sheets, including *United States v. International Paper Co.*, No. H-78-11 and *United States v. Boise Cascade Corp.*, No. H-78-12;
 - v. *In re Corrugated Container Antitrust Litigation;*
 - vi. *In the Matter of Stone Container Corp.*, FTC Docket No. C-8306; and
 - vii. *In re Linerboard Antitrust Litigation.*

REQUEST NO. 68:

All documents relating to compliance policies whether implemented, adopted, used or considered concerning federal, state, or international antitrust laws, competition laws, anti-monopoly laws, anti-cartel laws, unfair competition laws, anti-bribery laws and any similar laws or regulations, or contacts and/or communications with your competitors (including current and former versions of compliance policies and procedures, presentations, seminars, programs, memos, Corporate Communications, statements signed by your employees with responsibility for Containerboard Products that acknowledge receipt of or compliance with such policies arising from or relating to any prior legal proceedings) and all documents relating to any inquiries or investigations concerning compliance with such policies including those listed in the prior Request.

REQUEST NO. 69:

All documents relating to any joint defense, contribution, indemnity or judgment sharing agreement relating to any investigation, civil or criminal litigation involving the production, sale, resale or marketing of Containerboard Products.

Plaintiffs

REQUEST NO. 70:

All documents relating to the Named Plaintiffs in this action.

Complaints

REQUEST NO. 71:

All documents relating to complaints received from any Containerboard Products customer regarding prices, pricing, or terms or conditions of sale or the refusal or failure to

supply Containerboard Products and any responses thereto, and all documents reflecting any policy of your company concerning the handling of such complaints.

Affirmative Defenses

REQUEST NO. 72:

All documents that support any affirmative defense relating to the applicable statute of limitation period in this case (including actual or inquiry notice of the claims alleged in the Complaint).

REQUEST NO. 73:

All documents that support any affirmative defense relating to any release of claims alleged in this lawsuit by any Named Plaintiff or member of the proposed class.

REQUEST NO. 74:

All documents that support any affirmative defense relating in whole or in part, to the doctrines of estoppel and laches.

REQUEST NO. 75:

All documents that support any other affirmative defense you have alleged in your answer to the Complaint.

Forms of Contract

REQUEST NO. 76:

All forms of contracts or invoices issued to customers for sale of Containerboard products in the U.S.

ESI

REQUEST NO. 77:

Documents sufficient to identify and describe your computer hardware and systems, software, ESI, database, storage, backup and archiving systems and communications systems and devises used in connection with any of the Corporate Functions.

REQUEST NO. 78:

All documents necessary to understand the operation of any of the computer hardware and systems, software, ESI, database, storage, backup and archiving systems and communications systems and devises information requested herein (including, but not limited to, documents describing or defining the fields contained in any such database file naming conventions and standards; Help features or documentation; password, encryption, and other security protocols; diskette, CD, DVD, and other removable media labeling standards; email storage conventions, *e.g.*, limitations on mailbox sizes/storage locations; schedule and logs for storage; software and hardware upgrades (including patches) for the relevant time period (including who and what organization conducted such upgrades); and Backup tape rotation

REQUEST NO. 79:

Documents sufficient to identify and describe policies relating to your computer hardware and systems, software, ESI, database, storage, backup and archiving systems and communications systems and devises including, but not limited to, the following:

- a. employee usage policies for company computers, telephones, and other technology (including personal or home computer usage policies for work-related activities);
- b. company-wide monitoring software;

- c. Backup tape rotation schedules; and
- d. ESI retention, preservation and destruction schedules.

REQUEST NO. 80:

Documents sufficient to identify or describe the outputs of your ESI, information systems, databases and communications systems used by you to record, store, communicate, compute, analyze, organize, or retrieve information relating to the Corporate Functions (including all mainframe systems, legacy systems, archives, computers, networks, telephones, handheld devices, smart phones, storage or archiving systems, and word processing, electronic mail, personal information managers, calendar, or spreadsheet programs).

Data

REQUEST NO. 81:

All weekly, monthly, quarterly, and annual audited or unaudited financial statements, regularly prepared or received by your current or former directors, officers, Management, Planning Staff, Affiliated Person responsibilities relating to the Corporate Functions (including the CFO, treasurer or controller) for Containerboard Products sold in the U.S., including all documents appended to the foregoing (including profit and loss statements, balance sheets, cash flow statements, and other financial, accounting or operating statements) on a: (i) consolidated basis; (ii) for each business unit or segment; (iii) product line or product category basis; and (iv) facility-by-facility basis.

FOR THE FOLLOWING DOCUMENT REQUESTS NOS. 82 – 94 THE RELEVANT TIME PERIOD IS JANUARY 1, 2000 THROUGH DECEMBER 31, 2010 AND DATA REQUESTS ARE FOR ESI IN NATIVE FORMAT:

REQUEST NO. 82:

Documents sufficient to show your weekly, monthly, quarterly, and annual production of Containerboard Products.

REQUEST NO. 83:

Documents sufficient to show your weekly, monthly, quarterly, and annual price, cost and margin data for Containerboard Products sold in the U.S.

REQUEST NO. 84:

For all sales of Containerboard Products in the U.S., transaction level data, including:

- a. customer names, addresses and transaction dates;
- b. product data;
- c. price and volume data;
- d. discount, credit and rebate data;
- e. freight charge, transportation and other cost data; and
- f. any other data available in such a database.

REQUEST NO. 85:

Documents sufficient to show your weekly, monthly, quarterly, and annual sales (in dollars) and shipments (in short tons, linear feet, or other unit of volume) of Containerboard Products sold in the U.S. on a (i) consolidated basis; (ii) business segment or unit; (iii) product line or product category basis; and (iv) facility-by-facility basis.

REQUEST NO. 86:

Documents sufficient to show your weekly, monthly, quarterly, and annual imports of Containerboard Products sold in the U.S., identified by product line or product category⁹ and country of origination.

⁹ Includes brand, class, type, grade, or category

REQUEST NO. 87:

Documents sufficient to show your weekly, monthly, quarterly, and annual exports of Containerboard Products made in the U.S., identified by product line or product category and country of delivery or end-use.

REQUEST NO. 88:

Documents sufficient to show your weekly, monthly, quarterly, and annual inventories of Containerboard Products for sale in the U.S. identified on a (i) consolidated basis; (ii) business segment or unit; (iii) product line or product category basis; and (iv) facility-by-facility basis.

REQUEST NO. 89:

Documents sufficient to show your weekly, monthly, quarterly, and annual prices of Containerboard Products sold in the U.S. identified on a (i) consolidated basis; (ii) business segment or unit; (iii) product line or product category basis; and (iv) facility-by-facility basis..

REQUEST NO. 90:

For Containerboard Products sold in the U.S., all transaction level data for all costs of production (including raw materials), sales, marketing, and distribution, including:

- a. supplier names, addresses, and transaction dates;
- b. product data;
- c. price and volume data;
- d. discount, credit and rebate data;
- e. freight charge, transportation and other data; and
- f. any other data available in such a database.

REQUEST NO. 91:

For all Containerboard Products produced in the U.S., all transaction level data for all costs of production (including raw materials), sales, marketing, and distribution including:

- a. supplier names, addresses, and transaction dates;
- b. product data;
- c. price and volume data;
- d. discount, credit and rebate data;
- e. freight charge, transportation and other data; and
- f. any other data available in such a database.

REQUEST NO. 92:

Documents sufficient to show your weekly, monthly, quarterly, and annual cost incurred to produce Containerboard Products including, but not limited to, documents showing each of your fixed costs, variable costs (including raw materials, intermediates, labor, etc.), cost of goods sold, or any other measure of production cost, identified on a (i) consolidated basis; (ii) unit basis; (iii) product line or product category basis; and (iv) facility-by-facility basis. If any costs are allocated or shared costs, please produce all documents detailing the allocation and itemization of the costs so allocated or shared.

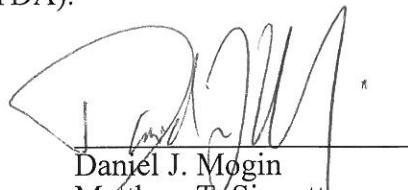
REQUEST NO. 93:

Documents sufficient to show your weekly, monthly, quarterly, and annual margins on Containerboard Products produced or sold in U.S., identified on a (i) consolidated basis; (ii) product line or product category basis; and (iii) facility-by-facility basis (including contribution margin, gross margin, other measure of margin).

REQUEST NO. 94:

To the extent not provided in response to the foregoing Request, documents sufficient to show your weekly, monthly, quarterly, and annual profits on Containerboard Products produced or sold in U.S reports identified on a (i) consolidated basis; (ii) product line or product category basis; and (iii) facility-by-facility basis, (including, but not limited to, gross profits, operating profits, net profits, cash flow reports, and EBITDA).

Dated: May 3, 2011



Daniel J. Mogin
Matthew T. Sinnott
THE MOGIN LAW FIRM, P.C.
707 Broadway, Suite 1000
San Diego, CA 92101
T: 619-687-6611
F: 619-687-6610

Michael J. Freed
Steven A. Kanner
FREED KANNER LONDON
& MILLEN LLC
2201 Waukegan Road, Suite 130
Bannockburn, IL 60015 USA
T: 224-632-4500
F: 224-632-4521
Lead Counsel for the Proposed Class

Exhibit A – Coding Field Labels for Document Discovery

Plaintiffs' Discovery Request	Coding Field Label
No. 1	Organizational
No. 2	Document Preservation
Nos. 3-6	Personnel
Nos. 7-11	Trade Associations
Nos. 12-17	Products
Nos. 18-20	Plants
Nos. 21-22	Production Capacity
Nos. 23-28	Capacity Changes
Nos. 29-37	Pricing
Nos. 38-44	Cost Accounting
Nos. 45-60	Competitive Conditions
No. 61	Competitor Contacts
No. 62	Industry Consolidation
No. 63	Corporate Communications
Nos. 64-69	Investigations
No. 70	Plaintiffs
No. 71	Complaints
Nos. 72-75	Affirmative Defenses
No. 76	Forms of Contract
Nos. 77-80	ESI – Systems

Plaintiffs' Discovery Request	Coding Field Label
No. 81	Data – Financial Statements
No. 82	ESI – Production
Nos. 83-84	ESI – Transaction Data
Nos. 85-86	ESI – Imports
No. 87	ESI – Exports
No. 88	ESI – Inventory
No. 89	ESI – Prices
Nos. 90-92	ESI – Costs
Nos. 93-94	ESI – Margins/Profits

CERTIFICATE OF SERVICE

I certify that on May 3, 2011, a copy of the foregoing Plaintiffs' First Request for Production of Documents Directed to All Defendants was served by email on the following persons:

SEE ATTACHED LIST

Daniel J. Mogin
Matthew T. Sinnott
THE MOGIN LAW FIRM, P.C.
707 Broadway, Suite 1000
San Diego, CA 92101
T: 619-687-6611
F: 619-687-6610

Kleen Products LLC, et al. v. Packaging Corporation of America, et al.,
No. 1:10-cv-05711

SERVICE LIST

Douglas J. Kurtenbach
Daniel E. Laytin
Barack S. Echols
Jeannie Young Evans
Luke Christian Ruse
KIRKLAND & ELLIS LLP
300 North LaSalle St
Chicago, IL 60654
(312) 682-2000 MAIN
(312) 826-3094 (LCR)
douglas.kurtenbach@kirkland.com
daniel.laytin@kirkland.com
bechols@kirkland.com
jeannie.evans@kirkland.com
juke.ruse@kirkland.com

Attorneys for Packaging Corp of America

Dao Lee Boyle
Scott Jared Fisher
NEAL, GERBER & EISENBERG
Two North LaSalle Street
Suite 1700
Chicago, IL 60602
(312) 269-8000 (DLB)
(312) 269-8035 (SJF)
dboyle@ngelaw.com
sfisher@ngelaw.com

Nathan P. Eimer
John K. Theis
EIMER STAHL KLEVORN & SOLBERG LLP
224 South Michigan Ave.
Suite 1100
Chicago, IL 60604
neimer@eimerstahl.com
jtheis@eimerstahl.com

Daniel W. Nelson
Joshua Lipton
Austin Schwing
GIBSON, DUNN & CRUTCHER LLP
1050 Connecticut Ave., N.W.
Washington, DC 20036
(202) 887-3687 MAIN
dnelson@gibsondunn.com
jlipton@gibsondunn.com

Kleen Products LLC, et al. v. Packaging Corporation of America, et al.,
No. 1:10-cv-05711

George A. Nicoud III
Casey N. Carrington
Rebecca Justice Lazarus
Robert Kim
Austin Schwing
Joel Willard
GIBSON, DUNN & CRUTCHER LLP
555 Mission St., Suite 3000
San Francisco, CA 94105
(415) 393-8308 MAIN
tnicoud@gibsondunn.com
ccarrington@gibsondunn.com
sdale@gibsondunn.com
rkim@gibsondunn.com
aschwing@gibsondunn.com

Attorneys for International Paper

Scott M. Mendel
John Edward Susoreny
Lauren Nicole Norris
K&L GATES LLP
70 West Madison, Suite 3100
Chicago, IL 60602
(312) 372-1121 (SMM)
(312) 807-4285 (JES)
(312) 807-4218 (LNN)
smendel@bellboyd.com
john.susoreny@klgates.com
lauren.norris@klgates.com

Attorneys for Cascades, Inc. and Norampac Industries

James R. Figliulo
Stephanie D. Jones
FIGLIULO & SILVERMAN
Ten South LaSalle Street
Suite 3600
Chicago, IL 60603
(312) 251-4600 MAIN
jfigliulo@fslegal.com
sjones@fslegal.com

Kleen Products LLC, et al. v. Packaging Corporation of America, et al.,
No. 1:10-cv-05711

Stephen Neuwirth
Sami Husayn Rashid
Kyle R. Taylor
QUINN EMANUEL URQUHART &
SULLIVAN LLP
51 Madison, 22nd FL
New York, NY 10010
(212)849-7000
stephenneuwirth@quinnmanuel.com
samirashid@quinnmanuel.com
kyletaylor@quinnmanuel.com

Attorneys for Georgia Pacific LLC

Andrew Stanley Marovitz
Britt Marie Miller
Courtney Lynn Anderson
MAYER BROWN LLP
71 South Wacker Drive
Chicago, IL 60606
(312) 782-0600 MAIN
amarovitz@mayerbrown.com
bmiller@mayerbrown.com
courtnofication@mayerbrown.com

Attorneys for Temple-Inland, Inc.

R. Mark McCareins
James Franklin Herbison
Michael P. Mayer
WINSTON & STRAWN LLP
35 West Wacker Drive
Chicago, IL 60601
(312) 558-5600 MAIN
rmccareins@winston.com
jherbison@winston.com
mmayer@winston.com

Attorneys for Smurfit-Stone Container Corp

David Marx Jr.
Jennifer A. Smulin Diver
Rachael V. Lewis
MCDERMOTT WILL & EMERY LLP
227 W. Monroe Street #4400
Chicago, Illinois 60606
(312) 984-7668
dmarx@mwe.com
jdiver@mwe.com
rlewis@mwe.com

Attorneys for Weyerhaeuser

Exhibit

B

**UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF ILLINOIS
EASTERN DIVISION**

KLEEN PRODUCTS LLC, *et al.*,

No. 1:10-cv-05711

Plaintiffs,

v.

PACKAGING CORPORATION OF
AMERICA, *et al.*,

Defendants.

PLAINTIFFS' NOTICE OF RULE 30(b)(6) DEPOSITIONS

Pursuant to Rule 30(b)(6) of the Federal Rules of Civil Procedure, Plaintiffs Kleen Products LLC; R.P.R. Enterprises, Inc.; Mighty Pac, Inc.; Ferraro Foods of North Carolina, LLC; Distributors Packaging Group, LLC; RHE Hatco, Inc.; and Chandler Packaging, Inc. will take the depositions of Defendants Packaging Corporation of America, International Paper, Cascades Canada, Inc., Norampac Holdings U.S., Inc., Weyerhauser Company, Georgia Pacific LLC, Temple-Inland Inc., and RockTenn CP, LLC, formerly known as Smurfit-Stone Container Corporation. The depositions will occur as follows:

Defendant	Date & Time	Location
Packaging Corporation of America	Monday, December 5, 2011, at 9:00 a.m. CST	Freed Kanner London & Millen LLC 2201 Waukegan Road, Suite 130 Bannockburn, IL 60015
International Paper	Tuesday, December 6, 2011, at 9:00 a.m. CST	Freed Kanner London & Millen LLC 2201 Waukegan Road, Suite 130 Bannockburn, IL 60015
Cascades Canada, Inc.	Wednesday, December 7, 2011, at 9:00 a.m. CST	Freed Kanner London & Millen LLC 2201 Waukegan Road, Suite 130 Bannockburn, IL 60015

Defendant	Date & Time	Location
Norampac Holdings U.S., Inc.	Thursday, December 8, 2011, at 9:00 a.m. CST	Freed Kanner London & Millen LLC 2201 Waukegan Road, Suite 130 Bannockburn, IL 60015
Weyerhauser Company	Friday, December 9, 2011, at 9:00 a.m. CST	Freed Kanner London & Millen LLC 2201 Waukegan Road, Suite 130 Bannockburn, IL 60015
Georgia Pacific LLC	Monday, December 12, 2011, at 9:00 a.m. CST	Freed Kanner London & Millen LLC 2201 Waukegan Road, Suite 130 Bannockburn, IL 60015
Temple-Inland Inc.	Tuesday, December 13, 2011, at 9:00 a.m. CST	Freed Kanner London & Millen LLC 2201 Waukegan Road, Suite 130 Bannockburn, IL 60015
RockTenn CP, LLC [f/k/a Smurfit-Stone Container Corporation]	Thursday, December 15, 2011, at 9:00 a.m. CST	Freed Kanner London & Millen LLC 2201 Waukegan Road, Suite 130 Bannockburn, IL 60015

Pursuant to Rule 30(b)(6), Defendants shall designate one or more witnesses to testify about the subject matters in the attached Schedule A. The depositions will be taken in accordance with and for all purposes permissible under the Federal Rules of Civil Procedure and will be videotaped and recorded by an officer duly qualified to administer oaths. All counsel of record are invited to attend and participate.

Dated: November 2nd, 2011



Steven A. Kanner
Michael J. Freed
FREED KANNER LONDON
& MILLEN LLC
2201 Waukegan Road, Suite 130
Bannockburn, IL 60015
T: 224-632-4500
F: 224-632-4521

Daniel J. Mogin
Matthew T. Sinnott
THE MOGIN LAW FIRM, P.C.
707 Broadway, Suite 1000
San Diego, CA 92101
T: 619-687-6611
F: 619-687-6610

Interim Co-Lead Counsel for the Proposed Class

30(b)(6) SCHEDULE A

DEFINITIONS

1. "And" and "or" are to be read interchangeably so as to give the broadest possible meaning to a particular request in which either or both is used.
2. The use of the singular form of any word includes the plural and vice versa, and the masculine, feminine and neuter form of any words includes each of the other genders.
3. The use of any tense of any verb shall also include within its meaning all other tenses of that verb.
4. "All" means "any and all" and the word "any" means "any and all."
5. "Communication" means, without limitation, the imparting or exchange of information, thoughts or opinions by any means including orally, in writing, by signs, signals or code, including oral, written or electronic communications, such as face-to-face meetings, electronic communications, emails, facsimiles, telephone communications, correspondence or other exchange of written or recorded information. The phrase "communication between" is defined to include instances where one party addresses the other party but the other party does not necessarily respond.
6. "Computer" means all devices utilizing microchips to facilitate processing, analysis, or storage of data, including microcomputers (also known as personal computers), laptop computers, portable computers, notebook computers, palmtop computers (also known as personal digital assistants or PDAs), minicomputers and mainframe computers.
7. "Database" shall include the term "data bank" and means and refer to any structured collection of electronic information organized into records or rows, together with all

other electronic data whose presence is required to analyze and view the information in a full and meaningful way.

8. "Defendant" means any company, organization, entity or person presently or subsequently named as a Defendant in this litigation including its predecessors, wholly-owned or controlled subsidiaries or affiliates, successors, parents, other subsidiaries, departments, divisions, joint ventures, other affiliates and any organization or entity managed or controlled by a named Defendant, including those merged with or acquired, together with all present and former directors, officers, employees, agents, attorneys, representatives or any persons acting or purporting to act on behalf of a Defendant.

9. "Document" shall be broadly interpreted as used in Rule 34(a) of the Federal Rules of Civil Procedure and means the original and all non-identical copies of all written, printed, recorded or electronically-stored information ("ESI"), including, but not limited to, letters, correspondence, memoranda, drafts, legal pleadings, calendars, diaries, day planners, journals, travel records, lists, outlines, summaries, notes, reports, compilations, notebooks, work papers, graphs, charts, spreadsheets, databases, blueprints, books, pamphlets, brochures, circulars, manuals, instructions, ledgers, drawings, sketches, telephone records, voicemail, telegrams, facsimiles, photographs, videotapes, audiotapes, digital audio or video recordings, other film and sound reproductions, presentations, slideshows, email, email attachments, text messages, instant messages, internal or external web sites, social networking data and messages, compact discs, computer files and disks, sales, advertising and promotional literature, agreements, stored recordings, minutes or other records of meetings, all written or graphic records or representations of any kind, and all electronic data, records or representations of any kind, including all file folders, directories, and metadata.

10. "Electronic Data" or "Data" means the original (or identical duplicate when the original is not available) and any non-identical copies (whether non-identical because of notes made on copies or attached comments, annotations, marks, transmission notations, or highlighting of any kind) of writings of every kind and description whether inscribed by mechanical, facsimile, electronic, magnetic, digital, or other means. Electronic Data includes, by way of example only, computer programs (whether private, commercial, or work-in-progress), programming notes or instructions, activity listings of electronic mail receipts and/or transmittals, output resulting from the use of any software program, including word processing documents, spreadsheets, Database files, charts, graphs and outlines, electronic mail, operating systems, source code of all types, peripheral drivers, PIF files, batch files, ASCII files, .pdf (portable document format) files, and any and all miscellaneous files and/or file fragments, regardless of the media on which they reside and regardless of whether said electronic data consists in an active file, deleted file or file fragment. Electronic Data includes any and all items stored on computer memories, hard disks, floppy disks, CD-ROMs, removable media such as Zip disks, Jaz cartridges, Bernoulli Boxes and their equivalent, magnetic tapes of all types, microfiche, punched cards, punched tape, computer chips, including, but not limited to EPROM, PROM, RAM and ROM, on or in any other vehicle for digital data storage and/or transmittal. The term Electronic Data also includes the file, folder tabs and/or containers and labels appended to, or associated with, any physical storage device associated with each original and/or copy.

11. "Electronic Media" means any magnetic or other storage media device used to record Electronic Data. Electronic Media devices may include computer memories, hard disks, floppy disks, CD-ROM, removable media such as Bernoulli Boxes and their equivalent, magnetic tapes of all types, microfiche, punched cards, punched tape, computer chips, including,

but not limited to EPROM, PROM, RAM and ROM, or on or in any other vehicle for digital data storage and/or transmittal.

12. "Electronically stored information" ("ESI") includes, without limitation, the following:

- a. activity listings of electronic mail receipts and/or transmittals;
- b. output resulting from the use of any software program, including, without limitation, word processing documents, spreadsheets, database files, charts, graphs and outlines, electronic mail, Blackberry Messenger (or similar program including mobile instant messengers, web browser-based messengers or social network messengers) or bulletin board programs, operating and backup systems, source code, PRF files, PRY files, batch files, ASCII files, and all miscellaneous media on which they reside and regardless of whether said electronic data exists in an active file, a backup file or system, a deleted file or system, or file fragment; and
- c. any and all items stored on computer memories, hard disks, floppy disks, CD-ROM, magnetic tape, microfiche, cloud data center(s) or in any other vehicle for digital data storage or transmittal, such as, but not limited to, desktop computers, servers and other network computers, backup tapes or systems, laptop computers, home or personal computers used for business purposes, a personal digital assistant, *e.g.*, Palm Pilot, R.I.M., Blackberry, iPhone or similar device (generically known as "smart phones" or "PDAs"), external storage devices (such as "keychain" drives) and file folder tabs, or containers and labels appended to, or relating to, any physical storage device associated with each original or copy of all documents requested herein.

13. "Identity" or "identify" means:

- a. when used with reference to a natural person, to state his or her full name, and if known, his or her present home address, present business address, present home and business telephone numbers, present or last known position and business affiliation and contact information;
- b. when used in reference to any entity, such as a partnership, joint venture, trust or corporation, to state the full legal name of such entity, each name under which such entity does business, the entity's street address, the entity's telephone number, the identity of the chief operating officer, manager, trustee or other principal representative and the identity of those persons employed by or otherwise acting for such entity who are known or are believed to possess the knowledge or information responsive to the document request and for which the entity was identified; and
- c. when used with reference to documents, to state specifically:
 - i. the type of documents involved (*e.g.*, letter, inter-office memorandum, etc.), together with information sufficient to enable a person to locate the document, such as its date, the name of any addressee, the name of any signer, the title or heading of the documents and its approximate number of pages; and
 - ii. the identity of the person last known to have possession of the document, together with the present or last known location of the document. If any document was, but is no longer, in your possession, custody or control, identify the document, state what disposition was made of it and the reason for such disposition. In lieu of identifying documents, copies may be provided.

14. “Including” is used to emphasize certain types of documents requested and should not be construed as limiting the request in any way. Including therefore means “including, but not limited to,” or “including without limitation.”

15. “Litigation” refers to the above-captioned action.

16. “Meeting” means, without limitation, any assembly, convocation, encounter, communication or contemporaneous presence (whether in person or via any electronic, computer-assisted, digital, analog, video or telephonic method of communication) of two or more persons for any purpose, whether planned or arranged, scheduled or not.

17. “Network” means any hardware and/or software combination that connects two or more Computers together and which allows the Computers to share and/or transfer Data between them. For the purposes of this definition, the connection between or among the Computers need not be either physical or direct, *i.e.*, wireless Networks, and sharing and/or transferring Data via indirect routes utilizing modems and phone company facilities. In addition, there need not be a central file or Data server nor a central Network operating system in place, *i.e.*, peer-to-peer Networks and Networks utilizing a mainframe host to facilitate Data transfer.

18. “Person” means, without limitation, any individual, corporation, partnership or any variation thereof (e.g., limited partnership, limited liability partnership), limited liability company, proprietorship, joint venture, association, group or other form of legal entity or business existing under the laws of the United States, any state or any foreign country.

19. “Relating to,” “referring to,” “regarding,” “with respect to” or “concerning” mean without limitation the following concepts: concerning, constituting, discussing, describing, reflecting, dealing with, pertaining to, analyzing, evaluating, estimating, constituting, studying, surveying, projecting, assessing, recording, summarizing, criticizing, reporting, commenting, or

otherwise involving, in whole or in part, directly or indirectly. Documents are considered relating to the subject matter whether they are viewed alone or in combination with other documents.

RELEVANT TIME PERIOD

Unless otherwise stated, the Relevant Time Period for this Notice is January 1, 2000 to the present.

DEPOSITION SUBJECT MATTER

Defendant shall designate and produce for deposition one or more of its officers, directors, managing agents, or other persons who are knowledgeable about and consent to testify on its behalf concerning the following subject matters:

1. Document retention, archiving and/or destruction policies and/or procedures used by Defendant during the Relevant Time Period for the Documents and Data relevant to any claim or defense or any facts in this Litigation, including policies and procedures applicable to the retention and/or destruction of Documents, Electronic Data or other ESI, including the Electronic Media by which such information is maintained, the Network(s) by which such information is maintained and the location of such Media, Networks or hard copy Documents.

2. Defendant's policy and procedures for the suspension of any Document retention, archiving and/or destruction policies and/or procedures, including the details of any suspension of such policies and/or procedures relating to any Documents or Data relevant to any claim or defense or any facts in this Litigation.

3. The identity of all past or present subsidiaries, affiliates, business divisions, departments, groups, units and employees of Defendant that have preserved and/or have collected Documents and/or Data concerning or related to any claim or defense or any facts of the Litigation.

4. The process, by which Defendant has identified, collected and/or preserved Documents and/or Data concerning or related to any claim or defense or any facts of the Litigation.

5. The method employed to identify, collect, and/or preserve from each custodian's Documents and or Data.

6. Organizational policies regarding Documents, Electronic Data and Database management, retention, deletion, sharing, production, and backup of electronic media responsive to the claims or defenses or any facts of the litigation.

7. The identity and description of the following:

- a. the type, location and usage of any Computer and/or Network systems currently or previously in use by Defendant;
- b. the type, details, and usage of Network servers and mainframes currently or previously in use by Defendant, including details regarding Network users;
- c. the type, details and usage of Computer workstation operating systems, Databases and Network operating systems currently or previously in use by Defendant;
- d. the type, details, and usage of software applications, server applications, and Network applications currently or previously in use by Defendant, including:
 - i. email applications;
 - ii. personal and office planner applications;
 - iii. word-processing and database applications;
 - iv. instant messaging applications;
 - v. applications and programs relating to production, inventory management, and distribution
 - vi. any enterprise resource planning applications; and
 - vii. server applications.

8. The identity of all Computers, Databases, Networks that have been searched for Documents and/or Data responsive to any claim or defense or the facts of the Litigation.

9. The physical location where backup media are stored or catalogued, including off-site locations.

10. The identity of person(s) responsible for conducting, maintaining, storing and cataloguing any backup tapes made.

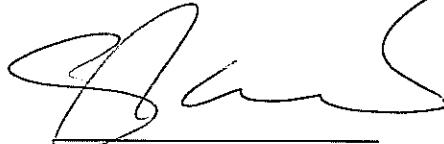
11. The identity of any backup media containing information responsive to any claim or defense or the facts of the Litigation, including where they are physically located, who maintains custodianship, dates of creation, content, and media type.

12. The identity of any backup media containing information responsive to any claim or defense or any facts of the Litigation that has been erased, copied over, misplaced, destroyed, or otherwise altered during the Relevant Time Period.

CERTIFICATE OF SERVICE

I certify that on November 24, 2011, a copy of the foregoing Plaintiffs' Notice of Rule 30(b)(6) Depositions was served by email on the following persons:

SEE ATTACHED LIST



Steven A. Kanner
Michael J. Freed
FREED KANNER LONDON
& MILLEN LLC
2201 Waukegan Road, Suite 130
Bannockburn, IL 60015
T: 224-632-4500
F: 224-632-4521

SERVICE LIST

Douglas J. Kurtenbach, P.C.
Daniel E. Laytin
Barack S. Echols
Leonid Feller
Jennie Y. Evans
Luke C. Ruse
Brenton Rogers
KIRKLAND & ELLIS LLP
300 North LaSalle
Chicago, IL 60654
(312) 862-2000
douglas.kurtenbach@kirkland.com
daniel.laytin@kirkland.com
barack.echols@kirkland.com
leonid.feller@kirkland.com
jeannie.evans@kirkland.com
luke.ruse@kirkland.com
brenton.rogers@kirkland.com

*Counsel for Defendant
Packaging Corporation of America*

Nathan P. Eimer
EIMER STAHL KLEVORN & SOLBERG
LLP
224 South Michigan Avenue, Suite 1100
Chicago, IL 60604
(312) 660-7600
neimer@eimerstahl.com

George A. Nicoud III
Austin V. Schwing
Daniel W. Nelson
Joshua Lipton
Casey N. Carrington
Robert E. Kim
Rebecca Justice Lazarus
Joel Willard
GIBSON, DUNN & CRUTCHER LLP
555 Mission Street
San Francisco, CA 94105
(415) 393-8200
tnicoud@gibsondunn.com
aschwing@gibsondunn.com
dnelson@gibsondunn.com
jlipton@gibsondunn.com
ccarrington@gibsondunn.com
rkim@gibsondunn.com
rjustice@gibsondunn.com
jwillard@gibsondunn.com

James T. McKeown
Trent M. Johnson
FOLEY & LARDNER LLP
777 East Wisconsin Ave.
Milwaukee, WI 53202
(414) 297-5530
jmckeown@foley.com
tjohnson@foley.com

Counsel for Defendant International Paper

Stephen Neuwirth
Sami H. Rashid
Kyle R. Taylor
QUINN EMANUEL URQUHART &
SULLIVAN, LLP
51 Madison Avenue, 22nd Floor
New York, NY 10010
(212) 849-7000
stephenneuwirth@quinnmanuel.com
samirashid@quinnmanuel.com
kyletaylor@quinnmanuel.com

R. Mark McCareins
Michael P. Mayer
James F. Herbison
WINSTON & STRAWN LLP
35 West Wacker Drive
Chicago, IL 60601
(312) 558-5600
rmccareins@winston.com
mmayer@winston.com
jherbison@winston.com

Counsel for Defendant RockTenn CP, LLC

James R. Figliulo
Stephanie D. Jones
FIGLIULO & SILVERMAN, P.C.
10 South LaSalle Street, Suite 3600
Chicago, IL 60603
(312) 251-4600
jfigliulo@fslegal.com
sjones@fslegal.com

Counsel for Defendant Georgia-Pacific LLC

Andrew S. Marovitz
Britt M. Miller
Courtney L. Anderson
Jaime Stilson
MAYER BROWN LLP
71 S. Wacker Drive
Chicago, IL 60606
(312) 782-0600
amarovitz@mayerbrown.com
bmiller@mayerbrown.com
clanderso@mayerbrown.com
jstilson@mayerbrown.com

Counsel for Defendant Temple-Inland Inc.

Scott M. Mendel
John E. Susoreny
Lauren N. Norris
K&L GATES LLP
70 W. MADISON ST.
SUITE 3100
CHICAGO, IL 60602
(312) 372-1121
scott.mendel@klgates.com
john.susoreny@klgates.com
lauren.norris@klgates.com

*Counsel for Defendants Cascades, Inc. and
Norampac Holdings U.S. Inc.*

David Marx Jr.
Jennifer S. Diver
Rachael V. Lewis
McDERMOTT WILL & EMERY LLP
227 W. Monroe Street
Chicago, IL 60606
(312) 372-2000
dmarx@mwe.com
jdiver@mwe.com
rlewis@mwe.com

*Counsel for Defendant Weyerhaeuser
Company*

Exhibit

C

DAVID D. LEWIS

David D. Lewis Consulting, LLC
1341 W. Fullerton Ave., #251
Chicago, IL 60614
Ph. (773) 975-0304
Fax (773) 289-0507
consult@DavidDLewis.com
www.DavidDLewis.com

Curriculum Vitae

Education & Awards

Fellow, American Association for the Advancement of Science (2006).

Ph.D. (1992) & M.S. (1988), Computer Science, Univ. of Massachusetts
• 1992 American Society for Information Science Doctoral Forum Award

B.S., Computer Science (1985) & B.A., Mathematics (1985), Michigan State Univ.
• Honeywell Futurist Award (1983)

Employment

Independent Consultant (2000-)

- Commercial & Nonprofit Clients: Designed algorithms, software, and processes for information retrieval, data mining, natural language processing, and statistical evaluation. Analyzed data sets. Trained personnel. Advised on competitors, market opportunities, and commercial products. Performed due diligence for investors. Co-wrote successful Phase I and Phase II SBIR grant proposals. Inventor on two patent applications. Consulting expert on three patent cases and one e-discovery case. Advised two investment groups on patent valuations. Served on scientific advisory boards.
- University-Based Research: Collaborated with university faculty, resulting in refereed publications and widely used open source software. Supervised students and developers. Managed subcontracts. Co-wrote research grants receiving funding of over 5 million dollars. Developed two major new information retrieval test collections, including first public test collection for e-discovery. Advised on technology transfer.

Co-Founder, Ornarose, Inc. (2001-2004)

- Co-founded startup to build text and data mining software. PI on two successful NSF SBIR Phase I grants.

Adjunct Assistant Professor (1998-2000), Univ. of Pennsylvania

- Presented guest lectures. Collaborated with U Penn faculty.

Principal Technical Staff Member (1996-2000), Technical Staff Member (1992-1996)

AT&T Bell Labs and AT&T Labs

- Designed and led development of text classification and data mining platform. Supervised summer students and developer. Conducted research leading to publications and five patents. Built prototypes for, and consulted with, AT&T business units. Advised AT&T and AT&T Ventures on acquisitions of external technology and companies. Developed new public test collections for information retrieval research.

Employment, cont.

Research Associate (rank: Assistant Professor) (1991-1992)

Center for Information and Language Studies, Univ. of Chicago

- Co-taught seminar on IR. Conducted research. Supervised research assistants and developers on IR and NLP projects. Wrote grant proposals and received two subcontracts totaling \$100,000. Developed and distributed first public test collection for text categorization research.

Independent Consultant (1988-1992)

- Consulting on information retrieval, natural language processing, and knowledge representation.

Research Assistant, Dept. of Computer Science, Univ. of Mass. (1985-1991)

- Conducted research on IR, NLP, and machine learning. Supervised developers and student employees. Prepared and presented class lectures. Wrote portions of grants receiving funding of \$2.4 million.

Summer Intern, Honeywell, Inc. (Summer 1984)

- Implemented NLP systems for relational database search and document processing.

Tutorials and Short Courses

- Information Retrieval for E-Discovery: *SIGIR '10* (accepted for presentation), OLP online course (accepted for presentation)
- Text Mining (published title varies): *JSM '06, Infornortics '07, SIGIR '07*.
- Introduction to Logistic Regression (published title varies): *SIGIR '05, KDD '05, AAAI '08, SIGIR '09*.
- Text Mining for Surveillance I. *DIMACS Tutorial on Statistical and Other Analytic Health Surveillance Methods*. (June 2003).
- Text Mining. *CCR/DIMACS Workshop on Mining Massive Data Sets and Streams: Mathematical Methods and Algorithms for Homeland Defense* (June 2002).
- Machine Learning for Information Organization (published title varies): *Infonortics '01, ICML '01, SIGIR '01, ASIS&T '01, ACL '02, SIGIR '02, ASIS&T '02, SIGIR '03, SIGIR '04, Infonortics '05*.
- Advanced Machine Learning for Information Retrieval (* w/ Yoram Singer): *SIGIR '99*, SIGIR '00**.
- Introduction to Machine Learning for Information Retrieval (* w/ Yoram Singer): *SIGIR '97, SIGIR '98, SIGIR '99*, SIGIR '00*, CIKM '00*.
- Natural Language Processing for Information Retrieval (* w/ Liz Liddy): *ACL '93*, SIGIR '93*, SIGIR '94*, ASIS '94*.

Organizational Activities

- Member of planning committees for DOD & NIST evaluations of NLP & IR systems (MUC-3 & MUC-4, TREC-1 through TREC-6; TREC Genomics Pre-Track Workshop (2002); TREC-2004 to present; TREC-2005, 2006, 2007 Genomics Track Steering Committees). Designed and directed the first and second NIST TREC text classification evaluation (TREC-4 and TREC-5 Filtering Tracks, 1996). Co-designed/directed with Jason Baron and Doug Oard the first NIST TREC electronic discovery evaluation (TREC-2006 Legal Track, 2006).
- Co-Organizer, *Workshop on Strategic Research Directions in AI: Knowledge Representation, Discovery, and Integration*, 2003.
- Chair & Founder, *Workshop on Operational Text Classification*, 2001 (OTC-01 & OTC-03). Organizing committee (OTC-02).
- Publicity Chair, *SIGIR '97*. Mentoring Chair (founded program), *SIGIR '01*. Awards Chair, *SIGIR '01*. Mentor, *SIGIR '03*. Best Paper Committee Chair, *SIGIR '03*.
- Session Leader, *NSF Information and Data Management Workshop*, 2000.
- Session Leader, *NSF IR Tools Workshop*, 1998.
- Founder and moderator (1995-), *DDLBETA (Text Classification Mailing List)*, 400+ subscribers.
- Secretary, *ACM Special Interest Group on Information Retrieval*, 1995-1997, 2003-2010.
- Program Chair (Information Retrieval), *Symposium on Document Analysis and Information Retrieval*, 1995.
- Co-chair, *AAAI Fall Symposium on Active Learning*, 1995.
- Co-chair, *Workshop on Learning in Intelligent Information Retrieval at Eighth International Workshop on Machine Learning*, 1991.

Other Service Activities

- External Advisory Board: Columbia Univ. Center for Research on Information Access (CRIA) (1996-?).
- Editorial boards: *ACM Transactions on Information Systems* (Associate Editor: 1994-2001), *Computational Linguistics* (1993-1995), *Encyclopedia of Information Retrieval* (1999-?), *Information Retrieval* (1997-2001), *Journal of Artificial Intelligence Research* (2000-2003), *Journal of Natural Language Engineering* (1994-2001).
- Co-Editor, *ACM Transactions on Information Systems Special Issue on Text Categorization* (1994).

- Program committees: SIGIR (1993-1998, 2000, 2002), SIGIR Best Paper Committee (2009), ARPA HLT (1993), WVLC (1993), ACL (1994), CIKM (1994), SDAIR (1994-1996), AAAI Workshop on Integrating Multiple Learned Models (1996), IJCAI Digital Libraries Workshop (1997), KDD (1997), AI & Statistics (1997-98), ICML (1998, 2000), Uncertainty (1999), WDA (2001), CEAS (2004), IJCNLP (2004), AIRWeb (2005), ECIR (2007), HLT (2009), ACL (2010), ECIR 2010 LSHC Workshop and others.
- Review panels: NSF SBIR (1993). NSF-CISE-IIS (2000).
- Graduate Student Representative, Computer Science Dept., U Mass Amherst (1987-1988).
- Cognitive Science Steering Committee, U Mass Amherst (1986-1988).
- Dissertation Committees: Michelle Keim (U. Washington, 1997). Ken Lang (CMU, unfinished: founder, *wisewire.com*). Marc Ringuette (CMU, unfinished: self-described software playboy). Aynur Dayanik (Rutgers, 2006).
- MentorNet mentor (1999-2001).

Publications and Talks

Patents (issued)

Chowdhury, A. R., Beitzel, S. M., Lewis, D. D., Kolcz, A. *Web query classification*. Patent No. 7,779,009 (August 17, 2010).

Guyon, I., Reiss, E. P., Doursat, R., Weston, J. A. E., Lewis, D. D. *Data mining platform for bioinformatics and other knowledge discovery*. Patent No. 7,542,947 (Jun 2, 2009).

Lewis, D. D., Singhal, A. K., and Stern, D. L. *Method and apparatus for communicating information in a supervised learning system*. Patent No. 6,523,017, Feb 18, 2003. Continuations: 6,668,248 (Dec 23, 2003); 6,931,383 (Aug 16, 2005).

Church, K. W., Heflman, J. I., and Lewis, D. D. *Method for retrieving texts which are similar to a sample text*. Patent No. 5,970,484, Oct 19, 1999.

Cohen, E. and Lewis, D. D. *Retrieval system and method*. Patent No. 5,950,189, Sep 7, 1999.

Knowles, K. A. and Lewis, D. D. *Finding an e-mail message to which another e-mail message is a response*. Patent No. 5,905,863, May 18, 1999.

Lewis, D. D. *Method and apparatus for training a text classifier*. Patent No. 5,675,710, Oct 7, 1997.

Catlett, J. A., Gale, W. A., and Lewis, D. D. *Training apparatus and method*. Patent No. 5,671,333, Sep 23, 1997.

Journal Articles

Oard, D. W., Baron, J. R., Hedin, B., Lewis, D. D., Tomlinson, S. (2010). Evaluation of Information Retrieval for E-Discovery. *Artificial Intelligence and Law*, 2010, to appear.

Beitzel, S. M., Jensen, E. C., Lewis, D. D., Chowdhury, A., Frieder, O. (2007). Automatic Classification of Web Queries Using Very Large Unlabeled Query Logs. *ACM Transactions on Information Systems*, vol 25, no. 2, April 2007.

Genkin, A., Lewis, D. D., and Madigan, D. (2007). Large-Scale Bayesian Logistic Regression for Text Categorization. *Technometrics*, vol. 49, no. 3, pp. 291-304, August 2007.

Lewis, D. D.; Yang, Y.; Rose, T.; and Li, F. RCV1: A New Benchmark Collection for Text Categorization Research. *Journal of Machine Learning Research*, vol. 5, pp. 361-397, 2004.

Cohen, E. and Lewis, D. D., Approximating Matrix Multiplication for Pattern Recognition Tasks. *Journal of Algorithms*, vol. 30, pp. 211-252, 1999.

Lewis, D. D. and Knowles, K. A., Threading Electronic Mail: A Preliminary Study. *Information Processing and Management*, vol. 33, pp. 209-217, 1997.

Lewis, D. D. and Sparck Jones, K., Natural Language Processing for Information Retrieval. *Communications of the ACM*, vol. 39, pp. 92-101, Jan, 1996.

Chinchor, N., Hirschman, L., and Lewis, D. D., Evaluating Message Understanding Systems: An Analysis of the Third Message Understanding Conference (MUC-3). *Computational Linguistics*, vol. 19, pp. 409-449, Sep, 1993.

Lewis, D. D., Croft, W. B., and Bhandaru, N., Language-Oriented Information Retrieval. *International Journal of Intelligent Systems*, vol. 4, pp. 285-318, 1989.

Refereed Conference Papers

Agam, G., Argamon, A., Frieder, O., Grossman, D., Lewis, D. Content-Based Document Image Retrieval in Complex Document Collections. *Document Recognition and Retrieval XIII, Proc. SPIE Int. Soc. Opt. Eng. 6500*, pp. 65000S-1 to 65000S-12, 2007.

Dayanik, A., Lewis, D., Madigan, D., Menkov, V., Genkin, A. Constructing Informative Prior Distributions from Domain Knowledge in Text Classification. *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 493-500, 2006.

Agam, G., Argamon, A., Frieder, O., Grossman, D., Lewis, D., Sohn, G., Voorhees, K. A Complex Document Information Processing Prototype. *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2006. (Poster paper.)

Lewis, D., Agam, G., Argamon, A., Frieder, O., Grossman, D., Heard, J. Building a Test Collection for Complex Document Information Processing. *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2006. (Poster paper.)

Lewis, D., Agam, G., Argamon, A., Frieder, O., Grossman, D., Heard, J. Building a Test Collection for Complex Document Information Processing. *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2006. (Poster paper.)

Beitzel, S., Jensen, E. C., Frieder, O., Grossman, D. Lewis, D. D., Chowdhury, A., Kolcz, A. Improving Automatic Web Query Classification via Semi-Supervised Learning. *ICDM '05: The Fifth IEEE International Conference on Data Mining*, 2005.

Agam, G., Argamon, S., Frieder, O., Grossman, D., Lewis, D. "Document Information Processing: Towards Software and Test Collections," *2005 Symposium on Document Image Understanding Technology*, 2005.

Beitzel, S., Jensen, E. C., Frieder, O., Grossman, D. Lewis, D. D., Chowdhury, A., Kolcz, A. "Automatic web query classification using labeled and unlabeled training data," *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 581-582, 2005. (Poster paper.)

Madigan, D., Genkin, A., Lewis, D. D., Fradkin, D. "Bayesian Multinomial Logistic Regression for Author Identification," *25th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, pp. 509-516, 2005.

Madigan, D., Genkin, A., Lewis, D. D., Argamon, S., Fradkin, D., Ye, L. "Author Identification on the Large Scale," *Joint Annual Meeting of the Interface and the Classification Society of North America*, 2005.

Madigan, D., Eyheramendy, S., Lewis, D. D., "On the Naive Bayes Model for Text Categorization," *Proceedings of the Ninth International Workshop on Artificial Intelligence & Statistics*, 2003.

Iyer, R. D., Lewis, D. D., Schapire, R. E., Singer, Y., and Singhal, A., "Boosting for Document Routing , " *Proceedings of the Ninth International Conference on Information and Knowledge Management*, McLean, VA, pp. 70-77, 2000.

Lewis, D. D., Stern, D. L., and Singhal, A., "ATTICS: A Software Platform for Online Text Classification," *Proceedings of SIGIR '99: 22nd International Conference on Research and Development in Information Retrieval*, pp. 267-268, 1999.

Keim, M., Lewis, D. D., and Madigan, D., "Bayesian Information Retrieval: Preliminary Evaluation," *Preliminary Papers of the Sixth International Workshop on Artificial Intelligence and Statistics*, pp. 303-310, 1997.

Cohen, E. and Lewis, D. D., "Approximating Matrix Multiplication for Pattern Recognition Tasks," *Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 682-691, 1997.

Lewis, D. D., Schapire, R. E., Callan, J. P., and Papka, R., "Training Algorithms for Linear Text Classifiers," *SIGIR '96: Proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 298-306, 1996.

Lewis, D. D., "Evaluating and Optimizing Autonomous Text Classification Systems," *SIGIR '95: Proceedings of the 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 246-254, 1995.

Ittner, D. J., Lewis, D. D., and Ahn, D. D., "Text Categorization of Low Quality Images," *Symposium on Document Analysis and Information Retrieval*, pp. 301-315, 1995.

Lewis, D. D. and Ringuelette, M., "A Comparison of Two Learning Algorithms for Text Categorization," *Third Annual Symposium on Document Analysis and Information Retrieval*, Las Vegas, NV, pp. 81-93, 1994.

Lewis, D. D. and Gale, W. A., "A Sequential Algorithm for Training Text Classifiers," *SIGIR 94: Proceedings of the Seventeenth Annual International ACM-SIGIR Conference on Research and Development in Information Retrieval*, pp. 3-12, 1994.

Lewis, D. D. and Catlett, J., "Heterogeneous Uncertainty Sampling for Supervised Learning," *Machine Learning: Proceedings of the Eleventh International Conference on Machine Learning*, pp. 148-156, 1994.

Church, K., Gale, W., Heflman, J., and Lewis, D., "Fax: An Alternative to SGML," *Proceedings of the 16th International Conference on Computational Linguistics (COLING '94)*, pp. 525-529, 1994.

Lewis, D. D., "An Evaluation of Phrasal and Clustered Representations on a Text Categorization Task," *Fifteenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 37-50, 1992.

Croft, W. B., Turtle, H. R., and Lewis, D. D., "The Use of Phrases and Structured Queries in Information Retrieval," *Fourteenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 32-45, 1991.

Lewis, D. D. and Croft, W. B., "Term Clustering of Syntactic Phrases," *Thirteenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 385-404, 1990.

Croft, W. B. and Lewis, D. D., "An Approach to Natural Language Processing for Document Retrieval," *Tenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 26-32, 1987.

Unrefereed Conference Proceedings, Book Chapters, and Miscellaneous Publications

Oard, D. W., Baron, J. R., Lewis, D. D. "Some Lessons Learned to Date from the TREC Legal Track (2006-2009)". Manuscript, Feb. 24, 2010. <http://trec-legal.umiacs.umd.edu/LessonsLearned.pdf>

Martentette, B. J., Bell, S. T., Adair, C. K., Fisher, D., Lewis, D. D., Gerding, D. "Team Process Measurement: Comparing Team Member and Observer Ratings," 25th Annual Conference of the Society for Industrial & Organizational Psychology (SIOP 2010), 2010. (Poster abstract in conference program.)

Agam, G., Argamon, S., Frieder, O., Grossman, D., Lewis, D. "Content-Based Document Image Retrieval in Complex Document Collections," *Document Recognition and Retrieval XIV*, Proc. SPIE Vol. 6500, pp. 65000S-1 to 65000S-12, 2007.

Baron, J. R., Lewis, D. D., Oard, D. W. "TREC-2006 Legal Track Overview," *Text REtrieval Conference (TREC 2006)*, 2007, pp. 79-98.

Agam, G., Argamon, S., Frieder, O., Grossman, D., Lewis, D. "Complex Document Information Processing: Prototype, Test Collection, and Evaluation," *Document Recognition and Retrieval XIII*, Proc. SPIE Vol. 6067, pp. 60670N-1 to 60670N-11, 2006.

Dayanik, A., Genkin, A., Kantor, P., Lewis, D. D., Madigan, D., Menkov, V. "DIMACS at the TREC 2005 Genomics Track," *Text REtrieval Conference (TREC 2005)*, 2006.

Agam, G., Argamon, S., Frieder, O., Grossman, D., Lewis, D. "Complex Document Information Processing: Towards Software and Test Collections," *First International Conference on Intelligence Analysis Methods and Tools*, 2005.

Dayanik, A., Fradkin, D., Genkin, A., Kantor, P., Lewis, D. D., Madigan, D., Menkov, V. "DIMACS at the TREC 2004 Genomics Track," *Text REtrieval Conference (TREC 2004)*, 2005.

Gomes, C., et al. Report of the Workshop on Strategic Research Directions in AI: Knowledge Representation, Discovery, and Integration. Lewis, D. (editor). Cornell University, 2003.

Angelescu, A., Boros, E., Lewis, D., Menkov, V., Neu, D., Kantor, P. "Rutgers Filtering Working at TREC 2002: Adaptive and Batch," *Text REtrieval Conference (TREC 2002)*, 2003.

Dumais, S. T., Sebastiani, F., and Lewis, D. D. "Report on the Workshop on Operational Text Classification Systems (OTC-02)," *SIGIR Forum*, vol. 36, no. 2, 2002.

Lewis, D. D., Pollak, T., and Romeo, S. "Supervised Learning for Classifying Nonprofit Programs: A Preliminary Study (Abstract)," *Workshop on Operational Text Classification Systems (OTC-02)*, 2002.

Allan, J., et al. "Challenges in Information Retrieval and Language Modeling: Report of a Workshop Held at the Center for Intelligent Information Retrieval, University of Massachusetts Amherst, September 2002," 2002.

Lewis, D. D. and Sebastiani, F., "Report on the Workshop on Operational Text Classification Systems (OTC-01)," *SIGIR Forum*, vol. 35, no. 2, 2001.

Lewis, D. D., "Applying Support Vector Machines to the TREC-2001 Batch Filtering and Routing Tasks," *Text REtrieval Conference (TREC 2001)*, 2002.

Lewis, D. D., "Machine Learning for Text Categorization: Background and Characteristics," *Proceedings of the Twenty-First National Online Meeting*, pp. 221-226, 2000.

Singhal, A., Choi, J., Hindle, D., Lewis, D. D., and Pereira, F., "AT&T at TREC-7," *The Seventh Text REtrieval Conference (TREC-7)*, pp. 239-252, 1999.

Condliff, M. K., Lewis, D. D., Madigan, D., and Posse, C., "Bayesian Mixed-Effects Models for Recommender Systems," *ACM SIGIR '99 Workshop on Recommender Systems: Algorithms and Evaluation*, 1999.

Lewis, D. D., "Naive (Bayes) at Forty: The Independence Assumption in Information Retrieval," *European Conference on Machine Learning*, pp. 4-15, 1998.

Lewis, D. D., "The TREC-5 Filtering Track," *The Fifth Text REtrieval Conference (TREC-5)*, pp. 76-96, 1997.

Lewis, D. D., "Information Retrieval and the Statistics of Large Data Sets," *Proceedings of the NRC Massive Data Sets Workshop*, pp. 143-147, 1997.

Cohn, D. and Lewis, D., Active Learning Symposium *AI Magazine*, vol. 17, pp. 83Spring, 1996.

Lewis, D. D., "The TREC-4 Filtering Track," *The Fourth Text REtrieval Conference (TREC-4)*, pp. 165-180, 1996.

Lewis, D. D., "Challenges in Machine Learning for Text Classification," *Proceedings of the Ninth Annual Conference on Computational Learning Theory*, pp. 1, 1996.

Lewis, D. D., "Active by Accident: Relevance Feedback in Information Retrieval," *AAAI Fall Symposium on Active Learning*, Nov. 1995.

Lewis, D. D., A Sequential Algorithm for Training Text Classifiers: Corrigendum and Additional Data *SIGIR Forum*, vol. 29, pp. 13-19, Fall, 1995.

Kwok, K. L., Grunfeld, L., and Lewis, D. D., "TREC-3 Ad-Hoc, Routing Retrieval and Thresholding Experiments Using PIRCS," *Overview of the Third Text REtrieval Conference (TREC-3)*, pp. 247-255, Apr. 1995.

Lewis, D. D., "A Brief Overview of Information Retrieval," *Proceedings IEEE Automatic Speech Recognition Workshop*. Snowbird, Utah, pp. 66, 1995.

Lewis, D. D. and Hayes, P. J., Guest Editorial *ACM Transactions on Information Systems*, vol. 12, pp. 231Jul, 1994.

Callan, J. P. and Lewis, D. D., "The Efficiency Issues Workshop Report," *The Second Text Retrieval Conference (TREC-2)*, pp. 303-304, Mar. 1994.

Lewis, D. D., "Natural Language Processing and Text Classification: Position Paper," *Proceedings of the Workshop on Future Directions in Text Analysis, Retrieval, and Understanding*, Chicago, IL, pp. 52-57, Feb. 1993.

Lewis, D. and Smeaton, A. "Workshop on: Use of Natural Language Processing at TREC," The First Text REtrieval Conference (TREC-1), pp. 365-366, 1993.

Lewis, D. D., Representation and Learning in Information Retrieval 1992. Department of Computer Science; Univ. of Massachusetts. Ph.D. Thesis.

Lewis, D. D. and Tong, R. M., "Text Filtering in MUC-3 and MUC-4," *Proceedings of the Fourth Message Understanding Conference (MUC-4)*, McLean, VA, pp. 51-66, 1992.

Lewis, D. D. Text Representation for Intelligent Text Retrieval: A Classification-Oriented View. In: *Text-Based Intelligent Systems*, ed. Jacobs, P. S. Hillsdale, NJ: Lawrence Erlbaum, 1992. pp. 179-197.

Lewis, D. D., "Feature Selection and Feature Extraction for Text Categorization," *Proceedings of Speech and Natural Language Workshop*, Harriman, NY, 1992.

Lewis, D. D., "Learning in Intelligent Information Retrieval," *Eighth International Workshop on Machine Learning*, pp. 235-239, 1991.

Lewis, D. D., "Evaluating Text Categorization," *Proceedings of Speech and Natural Language Workshop*, pp. 312-318, 1991.

Lewis, D. D., "Data Extraction as Text Categorization: An Experiment with the MUC-3 Corpus," *Proceedings of the Third Message Understanding Evaluation and Conference*, San Diego, CA, 1991.

Lewis, D. D., "Representation Quality in Text Retrieval: An Introduction and Experiment," *Proceedings of Speech and Natural Language Workshop*, pp. 288-295, 1990.

Lewis, David D. Desiderata for IR Parsing, and an Evaluation of the Transportable Understanding Mechanism Package (TRUMP). COINS, U Mass Amherst. M.S. Thesis.

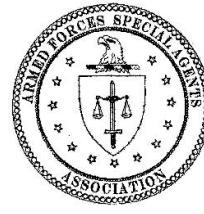
Invited Talks

- (Naive) Bayesian Text Classification for Spam Filtering. *Minimizing Market MAYhem: Statistical Applications in Direct Marketing, the Spring Conference of the Chicago Chapter of the American Statistical Association*, 2004.
- Text Mining: New Technologies from Old Concepts. *NFAIS*, 2004. (Panel)
- (Spam vs.) Forty Years of Machine Learning for Text Classification. *Spam Conference*, 2003.
- That Sounds Like "Mine": Some Thoughts on Text and Speech Mining. *AT&T Speech Days* (Oct. 2002)
- Text Mining. *Seventh Annual Text Summit* (Thomson Corporation internal conference - Sept. 2002)
- Generative and Discriminative Models in Text Classification. *Workshop on Challenges in Information Retrieval and Language Modeling* (Sept. 2002)
- Automatic Text Categorization. Testimony for hearing of National Research Council project on *Tools and Strategies for Protecting Kids from Pornography and Their Applicability to Other Inappropriate Internet Content* (Mar. 2001).
- Classifying and Mining Documents. *SPIE Document Recognition and Retrieval IX* (Jan. 2001).
- Next-Generation Information Access on the Web (panel). *ACM Conference on Information and Knowledge Management* (Nov. 2000).
- Text Categorization and Text Mining, *ACM SIGKDD International Conference on Knowledge Discovery & Data Mining: Workshop on Text Mining*, (Aug 2000).
- Machine Learning for Text Categorization: Background and Characteristics, *National Online Meeting and IOLS*, (May 2000).
- Naive (Bayes) at Forty: The Independence Assumption in Information Retrieval, *European Conference on Machine Learning*, 1998.
- Research Directions in Information Retrieval, *NSF Information and Data Management Workshop*, 1998.
- What Have We Learned, and Where are We Going?, *Infonortics Search Engine Meeting*, 1998.
- Challenges in Machine Learning for Text Classification, *Conference on Computational Learning Theory*, 1996.

- An Overview of Information Retrieval, *IEEE Signal Processing Society Workshop on Automatic Speech Recognition*, 1995.
- An Industry View on Education for IR (panel), *ACM SIGIR Conference on Research and Development in Information Retrieval*, 1995.
- Language Processing in Academia and Industry (panel), *American Society for Information Science Annual Conference*, 1994.
- Statistical Language Processing in Information Retrieval (panel), *IEEE Conference on Artificial Intelligence for Applications*, 1993.
- Optical Character Recognition and IR (panel), *Symposium on Document Analysis and Information Retrieval*, 1993.
- Natural Language and Information Retrieval (panel), *Workshop on Future Directions in Text Analysis, Retrieval, and Understanding*, 1991.
- Language-Oriented Information Retrieval (panel), *American Society for Information Science Annual Conference*, 1988.
- Evaluation in Information Retrieval, *DARPA Workshop on the Evaluation of Natural Language Processing Systems*, 1988.
- Numerous invited talks at institutions, including AT&T (2), BBN (2), Bellcore, Brooklyn Poly, Cancer Research UK, Carnegie Mellon (4), CDC, CNI-IRI Pisa (2), Columbia (2), CUNY - Queens, Colorado State (2), Cornell (2), DEC, DePaul U., Fuji Xerox, General Electric, IBM, Johns Hopkins, Just Systems, Lincoln Labs, MathSoft, Michigan State, Microsoft –Cambridge, Microsoft – Redmond (2), MIT (2), MITRE (2), National Library of Medicine, Naval Research Laboratory, NEC, NYU, Siemens, Synopsys, Syracuse U., Thinking Machines (2), Thomson Publishing, Toyota Technological Institute, Unisys, U. California Berkeley, U. Chicago, U. Dortmund, U. Glasgow, U. Illinois Chicago, U. Illinois UC, U Limerick, U. Massachusetts (2), U. Michigan, U. Pitt, U Maryland BC, U. Sheffield, U. Utah, U. Washington (3), U. Wisconsin (2), West Publishing, Xerox (2).

Exhibit

D



TIMOTHY D. HANNERS

1st Forensic Consulting LLC

Phone : (469) 774-6771

Computer Forensics Education

1994 - Phase IV Data Base Functional Users Course, Department of Defense, United States Air Force, 3300 Technical Training Wing – Computer Programming and System Maintenance

1993/1994 Basic and Advance Scientific Content Analysis – Behavioral Analysis in Written Statement

1989 – Air University/Community College of the US Air Force – Criminal Justice

1989 - US Air Force OSI Academy – Investigation Academy Investigators Course 89B

1989 - Certified Evidence Custodian, Department of Defense, United States Air Force Office of Special Investigations,

1990/1991- Child Sexual Abuse Investigations – Department of Defense, United States Air Force Office of Special Investigations Regional/Headquarters Training Courses, Forensic Instructors Leslie Bryant and Dawn Vincent

1993 - Child Sexual Abuse Tarrant County Community College

1994 – American Academy – Basic and Advance UNIX

1993 – US Air Force OSI Academy - Advance Investigators Course

1993 – US Air Force OSI Academy - Advance Surveillance Course

1994 – US Air Force OSI Academy - Computer Forensic Course, Taught by Howard Schmidt

1997 – Federal Law Enforcement Training Center – “Advance Telecommunications Fraud Investigations Training Program” and “Criminal Investigations in an Automated Environment” training programs (Basic and Advance Computer Crime Investigations including Computer Analysis, Chats, Hacking and Documentation of Findings)

1993/1994 – “Scientific Content Analysis”, Basic and Advanced Training with Laboratory for Scientific Interrogation.

1998 - Counterintelligence Force Protection Operations Workshop, Department of Defense, United States Air Force, Office of Special Investigations,

2000 - Ernst and Young, Advanced Forensics Course, sponsored by Ernst and Young, May 2000

2001 - System Forensics, Investigations, and Risk Management and Incident Response, sponsored by The SANS Institute – Online Investigations

2001 - Advanced Network Surveillance using Sniffer Pro, sponsored by Network Associates – Online Intercepts

2004 - Computer Forensics and Network Investigations, Vogen International Inc. Data Extraction, Examination, Keyword Searching, Advance Computer Collections, Unicode, Chat Decoding and Reporting.

2009 – ACE Certification from Access Data

I have approximately 140 credit hours in the Aviation Maintenance and Criminal Justice programs and obtained a Masters Certification in both Aviation Maintenance and Law Enforcement by the US Air Force.

Courses Taught Related to Computer Forensics

1994-1999 – DEPARTMENT OF DEFENSE, JUDGE ADVOCATE GENERAL (JAG), COMPUTER CRIME CONTINUING EDUCATION COURSE. HOW COMPUTERS RELATE TO CRIMINAL ACTIVITY AND HOW COMPUTER EVIDENCE SUPPORT THE ELEMENTS OF PROOF.

2000 – EVIDENCE COLLECTIONS IN AN AUTOMATED ENVIRONMENT. SEARCH AND SEIZURE TAUGHT TO PRIVATE SECTOR COMPANIES

2001 - LOCKHEED MARTIN, FORT WORTH, TX – LAW ENFORCEMENT ONLY, "ADVANCE COMPUTER FORENSIC FOR BANKING COMPUTERS"

2004 -2006 – ADVANCE COMPUTER FORENSIC ANALYSIS COURSE TAUGHT TO UNITED STATES SECRET SERVICE, US DEPARTMENT OF TREASURY INSPECTOR GENERAL'S OFFICE, ROYAL CANADIAN MOUNTED POLICE, SANS INSTITUTE ANNUAL CONFERENCE IN WASHINGTON DC.

2006 – COMPUTER FORENSIC BASIC INFORMATION TO KNOW IN CRIMINAL AND CIVIL DISCOVERY CASES, TO OKLAHOMA BAR ATTORNEY AND JUDGES CONFERENCE.

2007 – 2008 – COMPUTER FORENSIC DATA COLLECTION COURSE – TAUGHT IN DALLAS, TX; NEW YORK, NY; PHOENIX, AZ, DALLAS

2008 – DALLAS TEXAS BAR ASSOCIATION – CONTINUING LEGAL EDUCATION COURSE FOR COMPUTER INFORMATION IN CIVIL DISCOVERY ISSUES.

2007-2011 – INSTRUCTOR FOR COMPUTER FORENSIC BASIC COURSES AND ADVANCE COMPUTER FORENSICS COURSES – COURSE INCLUDED FORENSIC DOCUMENTATION, STORAGE, FORENSIC IMAGING OF COMPUTER SYSTEMS, LIVE COLLECTIONS AND EXACTIONS USING MAC WINDOWS AND LINUX OPERATING SYSTEMS.

Court Experience

Testified as an expert:

Arkansas VS Albert Keith Smith (CR2004-84-1) AR

Arkansas VS Deanna Bobo (G-CR-2005-168) AR

Lyon VS Lyon (Family Court SM43596) TX

Hib, Rogal & Hobbs VS Dave S. Derdeuff (CJ-05-9101-D) OK

United States of America VS Mark R. Emmons (Docket No. 08-0057) LA

Texas v. John Preston Creech (Case Number 401-82473-06) TX

John Doe VS Poudre School District (Case Number 2011CV1118) CO

Additionally have testified as an expert for the US Government in a Pennsylvania Parole Review Matter in 2003.

Additional Experience

I have written an opinion as an expert for a Special Master at the request of the 63rd Judicial District in Texas. The opinion addressed an issue before the court on how computers store information and the impact of random power surges on electronically stored media.

Conducted 3500+ computer related investigations/investigative leads over the last 16 years and have worked with computers for over 30 years.

Selected to assist with any voting machine issues for the 2008 presidential elections for the states of Florida, Texas and Arizona by the Republican National Committee (RNC).

PROFESSIONAL HISTORY

08/2011 - Present

President

1st Forensic Consulting LLC.

Lucas, Texas

- Serving as both owner and chief operations officer, I direct the company in every aspect.
- As a fully licensed Private Investigations Company we offer our years of experience and skills in the investigative field to providing the best in class services to our clients.
- Our experience includes Physical Security, Protective Services, Computer Forensics, Data Recovery and investigative services for both prosecution and Defense. We have worked large Civil Investigations involving millions of dollars.
- I develop new clients and manage existing client relationships.
- I oversee all aspects of the budgeting process, operations process, new hires, and facilities and equipment necessary to make the company function.
- I develop new techniques, equipment and processes necessary to make the company successful. This includes scientific techniques used in computer forensics, data recovery, investigative processes, evidence storage, evidence collections and techniques used in examinations conducted by the company.
- I conduct training exercises for law enforcement, legal service providers to include attorneys litigation support managers and paralegals.
- I provide expert witness services, interviewing services and investigative consultation services for my clients.

05/2007 - 08/2011

Director Forensic Services

Xact Data Discovery

Richardson, TX

- Direct all computer forensic activities for the firm.
- Manage the forensic division to ensure availability of all personnel and equipment to meet our mission objectives.
- Oversight of all stages in the Forensic process: from initial stages of client consultation, to forensic imaging, to performing analysis and preparing client reports and testifying As an expert.
- Manage tier one client relationships and assist our sales and production staff to meet goals for the company.
- Default new methods for conducting computer forensics which allowed an increase in our collections process which was unmatched in the industry.

04/2006 - 05/2007

President Eagle Forensic Services LLC

Eagle Forensic Services LLC

1600 Ports O' Call, Suite 100

Plano, TX 75075

- Provided professional computer forensic services including Computer Forensics Examinations and Analysis, Risk Management, Data Recovery, Discovery Services, Expert Testimony, Interviewing of Clients Employees, Physical Security, Personal Security, Vulnerability Assessments, Incident Response, Network Security, Intrusion Detection, Training to employees and attorneys, Network Surveillance, Attack and Penetration Testing for Clients.

02/2004 to 04/2006

Computer Investigations Manager (USA)

VOGON INTERNATIONAL LLC (US) (Purchased by Kroll Ontrack Inc April 2006)

RIVERSIDE CENTER

Suite 2625

2600 VAN BUREN

NORMAN, OK 73072

- Responsible for: Directing Computer Investigative Service offerings including Computer Forensics, Data Recovery, Laboratory Examinations, Expert Testimony, Interviews of Clients Employees, Physical Security, Personal Security, Vulnerability Assessments, Incident Response, Network Security, Intrusion Detection, Training, Network Surveillance, Attack and Penetration Testing for Clients on a Global Basis.
- Developed training processes for employees and clients and sold those services to the US Government and private sector companies resulting in additional revenues for the company.
- Responsible for client satisfaction resulting in significant repeat business including 3 million dollars in repeat business from one client in two years.
- Managed personnel assignments, marketing resources, and funding to meet company/client objectives.
- Responsible for testing, evaluation and submission of upgrade requirements of new and existing computer forensic equipment and forensic software.

10/2002 - 01/2004

Senior Computer Scientist

COMPUTER SCIENCES CORPORATION (CSC)

Corporate Headquarters

2100 East Grand Avenue

EL SEGUNDO, CA 90245

- As a Computer Forensic Analyst, managed computer forensics laboratory investigations, examination of computer hard drives, network log files and other storage media for evidence related to wrongful and/or illegal activities by computer users.
- Reported findings of examinations and provided testimony in US courts.
- Tested and evaluated computer software and hardware for use in examination of computer evidence.
- Provided consulting services for ongoing investigations and provided training to USPIS personnel pertaining to search and seizure of computer evidence.
- Managed contract relationship for Computer Sciences Corporation and was the primary point of contact for the contract.
- Conducted approximately 52 investigations for the postal inspection service.

01/2002 - 10/2002

Computer Forensic Principal

SECANT NETWORK

5068 W. Plano Parkway

Suite 300

Plano, TX 45093-4408

- Secant Networks was a startup company and owned by a business acquaintance.
- Assisted in the initial startup and worked to help build a service offering. Efforts resulted in the company growing and obtaining contracts with Fortune 500 companies.
- Duties included lecturing, marketing, and directing and performing forensic investigations.
- Investigations conducted for Secant Networks included both civil and criminal investigations.
- Conducted physical security assessments, assessments of video surveillance systems, security policies and risk management.

07/1999 - 01/2002

Vice President, Information Security
BANK OF AMERICA
Information Security
One Main Place
Dallas, TX 75202

- Directed and managed incident response to any computer related incident involving bank computers.
- Conducted analysis of events and procedures involving banking computers.
- Formulated plan to prevent wrongful activity and to minimize risk and exposure to the bank.
- Tested and evaluated new systems and software prior to implementation to ensure security requirements were addressed.
- Served as liaison to other corporate departments for security planning.

06/1978 - 07/1999

Special Operations/Special Agent/Computer Crime Investigator
UNITED STATES AIR FORCE OFFICE OF SPECIAL INVESTIGATIONS
Travis Air Force Base, CA 94535

- Collection, preservation, analysis, and presentation of computer-related evidence.
- Intercepted, identified and traced hostile activities directed towards US Government computer systems.
- Conducted computer related investigations of matters involving Foreign Intelligence Services.
- Conducted technical interviews in support of investigations regarding use of computers as tools of a crime.
- Performed investigations, protective service operations, and counter-intelligence operations.
- Responsible for defensive operations against techniques used by hostile intelligence services and subversive groups directed at or affecting the Air Force and national security.
- Served on details for protection of designated US and foreign dignitaries to include the President, Vice President, First Lady, Chairman of the Joint Chiefs of Staff and Secretary of Defense.
- Expert in counter-intelligence, force protection, defensive tactics, firearms and other weapons, surveillance and surveillance detection, anti-terrorism techniques, crime scene processing, computer forensics, forensic archeology, arson, interrogations and interviews.
- Served as Counter-intelligence Analyst supporting USAF Central Command in Southwest Asia 1998.
- Special Agent in Charge of Anti-Terrorism Mission Combine Forces (US, British and Canadian) in Central Africa in 1994 and 1996.
- While assigned to the USAF OSI, I conducted over 6,000 hours of direct analysis of computer systems for law enforcement and other national intelligence agencies.
- As a Special Agent, I have attended basic and advance investigations training which included documentation of computer records, crime scene reconstruction and Scientific Content Analysis of statements.

LICENSES, CERTIFICATIONS AND MEMBERSHIPS

Licenses

- Business license to practice within the state of Texas.
- Licensed private investigator currently within the state of Texas.

Certifications

- Access Data Certified Examiner (ACE)
- Certified in advanced computer evidence collection by the Federal Law Enforcement Training Center and Air Force Office of Special Investigations Academy, Ernest in Young's national laboratories, Vogon International,
- Additional training received from High-Tech Computer Crime Investigators Association HTcia.
- Certified in advanced telecommunications evidence collection by the Federal Law Enforcement Training Center.

- Certified in advanced computer forensics, physical security, forensic archeology, criminal and counter-intelligence investigations by Air Force Office of Special Investigations Academy, SANS Institute and Ernst and Young National Laboratory Inc.
- Advance training and application in counter-terrorism by US Air Force Central Command.

Memberships

- Member of Institute of Computer Forensic Professionals
- Member of Armed Forces Special Agents Association
- Member of Association of Former Air Force Office of Special Investigations Special Agents

Clearance History

- Top Secret/SBI/SCI and higher access: USAF – Last updated June 1998.
- Secret US Postal Inspection Service – Last updated February 2004.

Decorations, Medals, Badges, Citations and Campaign Ribbons

Meritorious Service Medal
Joint Service Commendation Medal
Air Force Commendation Medal with 3 Devices
Air Force Achievement Medal with 3 Devices
Armed Forces Expeditionary Medal with 1 Device
Humanitarian Service Medal
Air Force Good Conduct Medal with 5 Devices
Air Force Outstanding Unit Award with 1 Device
Air Force Organizational Excellence Award with 4 Devices
National Defense Service Medal
Air Force Overseas Long Tour Ribbon
Air Force Longevity Service Award Ribbon with 4 Devices
NCO PME Graduate Ribbon with 2 Devices
Small Arms Expert Marksmanship Ribbon with 1 Device
Air Force Training Ribbon
Numerous awards and letters recognizing performance and achievements

References

William L. Goode

The Goode Law Firm, A P.L.C.
812 Johnston Street
Lafayette, LA 70501
Phone: 337-234-0600
Fax: 337-237-1614
bill@goodelawyer.com

John Charles Hardin

John Charles Hardin Law Office
1650 West Virginia Street, #110
McKinney, TX 75069-4321
(972) 562-5291

Stephen G. Gleboff

Partner
Little Pedersen Fankhauser LLP
901 Main Street, Suite 4110
Dallas, Texas 75202
214.573.2300
214.573.2323 fax

Laura McConnell-Corbyn

Hartzog Conger Cason & Neville
1600 Bank of Oklahoma Plaza
201 Robert S. Kerr Avenue
Oklahoma City, OK 73102
405-235-7000 - Telephone
405-996-3403 - Facsimile
lmcconnell@hartzoglaw.com

Jeremy Tubb

Partner
Crowe & Dunleavy
20 North Broadway, Suite 1800
Oklahoma City, OK 73102
(405) 239-6681 - Direct Line
(405) 272-5267 - Direct Fax

Chris Brown

Associate
Winstead PC
5400 Renaissance Tower, 1201 Elm Street
Dallas, TX 75270
cabrown@winstead.com
214.745.5215

Exhibit

E

1

1214KDASC CONFERENCE
1 UNITED STATES DISTRICT COURT
1 SOUTHERN DISTRICT OF NEW YORK
2 -----x
2
3 MONIQUE Da SILVA MOORE, et
3 al.,
4
4 Plaintiffs,
5
5 v. 11 CV 1279 (RJS)
6
6 PUBLICIS GROUPE, et al.,
7
7 Defendants.
8
8 -----x

9 New York, N.Y.
9 January 4, 2011
10 10:58 a.m.

11 Before:

12 HON. ANDREW J. PECK,
12
13 Magistrate Judge

14 APPEARANCES
14

15 SANFORD WITTELS & HEISLER LLP
15 Attorneys for Plaintiffs
16 JANETTE WIPPER
16 DEEPIKA BAINS
17
17 JACKSON LEWIS LLP
18 Attorneys for Defendant MSL Group
18 BRETT M. ANDERS
19 VICTORIA WOODIN CHAVEY
20 ALSO PRESENT:
20

21 PAUL J. NEALE, DOAR Litigation Consulting
21 GENE KILMOV, DOAR Litigation Consulting
22 ERIC SEGGEBRUCH, Recomind
22 CRAIG CARPENTER, Recomind
23
24
25

SOUTHERN DISTRICT REPORTERS, P.C.
(212) 805-0300

1214KDASC

CONFERENCE

1 searches that apparently are going to be run across everybody's
2 email -- you're going to get a lot of stuff from so-called
3 comparators that isn't relevant, it doesn't make sense to do it
4 as a uniform group. If you want to say that you've got certain
5 comparators who you want different searches run on, that's a
6 different story. It doesn't make sense to pull all their
7 material in because they're a comparator at the same level.

8 MS. WIPPER: We would agree to that.

9 THE COURT: All right, my 12:00 o'clock call has
10 called in. I have lunch at 1:00. I'm hoping this won't take
11 long. Do you all want to just sit here or do you want to go
12 into the jury room and maybe work out some of these issues?
13 Go, lawyers and consultants, as needed, into the jury room. Do
14 not leave there. We will come get you after I deal with this
15 call.

16 (Recess)

17 THE COURT: OK, it's somewhere between 12:40 and
18 12:45. We're back on the record after my other conference.

19 What progress have you made? Or perhaps the other way
20 of looking at it is: What is it in the 15 minutes we have left
21 before lunch that you want me to rule on or give you advice on
22 with respect to the ESI protocol?

23 MR. ANDERS: Your Honor, we spent the bulk of the time
24 talking about the custodian list. We have identified five
25 custodians that are, I think, more on the either comparator

SOUTHERN DISTRICT REPORTERS, P.C.

(212) 805-0300

1214KDASC

CONFERENCE

1 category or secondary category where I think your Honor
2 suggested that maybe those email accounts get filtered prior to
3 being put into the database -- that's what we were trying to
4 understand -- but we have identified five where at least
5 plaintiffs would be willing to apply some type of keyword
6 search in the filtering to them first.

7 THE COURT: All right.

8 Ms. Wipper?

9 MS. WIPPER: With respect to the custodians, I believe
10 that the parties would be able to work it out. What we would
11 like to hear from the Court is your view on the differences
12 between the two protocols. Our protocol is --

13 THE COURT: I have no idea.

14 MS. WIPPER: OK.

15 THE COURT: When you send me 50 pages each, late at
16 night and/or the morning of, when you knew this conference was
17 scheduled for quite some time, there's a limit, and it was not
18 done as a redline or anything else as to where your differences
19 are. So you tell me what it would be most helpful for you, for
20 the ten minutes or so we have left, to rule on or advise on,
21 and I'll deal with it.

22 MR. ANDERS: Your Honor, I think that the key issue is
23 how we use predictive coding, and that's where there's
24 probably -- that's why we have our experts here, our vendors.

25 The way defendant MSL proposes using the predictive
SOUTHERN DISTRICT REPORTERS, P.C.
(212) 805-0300

1214KDASC

CONFERENCE

1 coding process would be as follows: We start with an initial
2 random sample, with a confidence level of 95 percent, with an
3 interval of plus or minus 2 percent. With the 3.2 million
4 document database, that random sample is 2,399 documents. We
5 have gone through those preliminarily. I had associates go
6 through those; I just finished going through it last night.

7 Of that 2,399 --

8 THE COURT: Just to stop you right there, my
9 understanding of predictive coding is that the coding, as
10 painful as it is, should be done by a very senior attorney,
11 meaning partner level or very senior associate, not the usual
12 team of umpteen lower associates with a lower billing rates.

13 MR. ANDERS: That's why I reviewed it, your Honor.

14 THE COURT: Well, as "reviewed it" as every one of the
15 coding decisions or spot-checked it?

16 MR. ANDERS: No, where I am right now is I have gone
17 through every one that was marked as relevant, I went through
18 400 so far that have been coded as not relevant, and I intend
19 to go through all of those but I first looked at the ones that
20 were relevant.

21 THE COURT: At the end of the process, you're going to
22 have done every single one of the --

23 MR. ANDERS: Yes.

24 THE COURT: Then I'm not sure why your client paid for
25 someone else to do it first, but that's not my problem, that's

SOUTHERN DISTRICT REPORTERS, P.C.

(212) 805-0300

1214KDASC

CONFERENCE

1 their problem.

2 OK, continue.

3 MR. ANDERS: So far 36 were deemed relevant. Of the
4 400 not relevant I have reviewed, they were clearly not
5 relevant. So right now the baseline is .015 percent of that
6 random sample was relevant. If you translate that to the
7 entire database, that's 48,000 documents.

8 After we did a random sample, then what we have done
9 at the same time is we have applied keywords and we have taken
10 the results of those keywords and sample-coded. So, for
11 example, if there's a keyword "reorganization," we may have
12 reviewed the top 200 random hits. We did that across the
13 board.

14 Also, to respond to several of plaintiffs' targeted
15 document requests, we ran targeted searches across the
16 database. That's what we have already produced, about a
17 thousand pages of documents. So we have that coding that's in
18 there.

19 Plaintiffs' counsel, they have sent us now three
20 different revisions of keywords. What I have proposed to
21 plaintiffs' counsel is, I'll give you the hit lists. I've
22 already given them two sets of hit lists; we have another set
23 to give them, I'll review -- or we'll review 3,000 of those
24 hits, you tell us how you want us to review it but pick the
25 hits, we'll review any of the top 200 in these ten categories,

SOUTHERN DISTRICT REPORTERS, P.C.

(212) 805-0300

1214KDASC

CONFERENCE

1 you tell us how to review it. We'll give them those results as
2 well.

3 Once that initial coding part is done, we'll let the
4 system go out, it will do a sample of, you know, train itself,
5 we'll get the results. Our proposal was to review, one, a
6 random sample of the results that come back as well as certain
7 judgmental sampling, share those results with plaintiff, they
8 can make their suggestions on how certain things should be
9 coded.

10 We have also identified six different categories that
11 documents can be coded towards. I think plaintiffs have asked
12 for us to do eight or nine. We can figure that out. Go
13 through that iterative process twice. At that point -- and
14 this is where sort of the proportionality and cost-limiting
15 comes in -- after we've gone through the iterative process
16 twice or if we have to go through another time, have the
17 computer give us the documents in rank order. And we have
18 agreed or proposed reviewing the top 40,000 rank documents.
19 And we arrived at that 40,000 document number -- we estimate it
20 will cost approximately \$200,000 using a five-dollar a document
21 cost estimate, it will cost 200,000 to review the 40,000.

22 When you take that 200,000 in review costs and you
23 couple it with our vendor costs, we're looking at a total spend
24 of approximately 550,000. We understand that plaintiffs take
25 issue with some of our vendor costs -- we can dispute that --

SOUTHERN DISTRICT REPORTERS, P.C.

(212) 805-0300

1214KDASC

CONFERENCE

1 but even just looking at the \$200,000 attorney fee review cost,
2 we think that that is a more than appropriate amount to spend
3 to see what we get. We have never told plaintiffs that we're
4 going to do this and this is all that you get. Our view is,
5 let's see what this yields us first, we think these are the
6 most relevant people, this is a sophisticated and excellent way
7 to find the cream of the crop, if you will. And after that
8 process is done, we'll be in a much better position to argue
9 and debate whether or not the incremental value of searching
10 another custodian is going to be worth the cost. And that's
11 essentially our view.

12 THE COURT: Let me hear from Ms. Wipper.

13 MS. WIPPER: Your Honor, we disagree with defense
14 counsel's position that the only issue is predictive coding,
15 because that kind of skips over a lot of other issues that --

16 THE COURT: Well, let's deal with the predictive
17 coding piece. I understand, from what little I have skimmed of
18 your proposal and theirs, that they're sort of only looking at
19 an email archive and you want lots of other steps looked at.

20 But assume that that other piece gets resolved,
21 meaning where they have to look, and maybe their 3.2 million
22 database will double or go up to whatever, but what's wrong
23 with the predictive coding methodology they have proposed,
24 which also sounds like it's being run on a fairly transparent
25 and cooperative basis?

SOUTHERN DISTRICT REPORTERS, P.C.

(212) 805-0300

1214KDASC

CONFERENCE

1 MS. WIPPER: Well, the main issue is cost because --
2 THE COURT: No, but where? In other words --
3 MS. WIPPER: It's impacting the methodology.
4 THE COURT: Well, the question becomes the review.
5 And my understanding of the way this works is by the time that
6 the system spits this out, and whether it's the top 40,000 or
7 whether the break point is 50,000 documents or 30,000, that
8 90-something percent of the relevant documents are going to be
9 found in the top hits, and that the costs of reviewing the rest
10 is not worth the candle in most cases.

11 Now, where that line gets drawn is something that I
12 can't decide until I've seen the results. In other words, when
13 one sees the results, as I understand it from this method, one
14 can see a sharp drop-off at a certain point, at which you then
15 still sample the documents that are not going to be reviewed,
16 and that's part of this whole iterative process.

17 If you are seeing that the top 40,000 documents give
18 you 90 percent of the responsive documents, and it's going to
19 cost a million dollars to go to the next hundred thousand
20 documents for eyes-on review, to get another 5 percent, it's
21 probably not worth it. If it's worth it to go to the top
22 50,000 because that's where the cliff line seems to be, that's
23 what people are going to have to do.

24 It also may be that once privilege is determined, that
25 they will let you -- the rest of this is so likely to be junk,

SOUTHERN DISTRICT REPORTERS, P.C.

(212) 805-0300

1214KDASC

CONFERENCE

1 that you want, under an attorneys'-eyes-only or some process,
2 an informal basis, you want to look at the documents that go
3 from 40,000 to 80,000, you can look at them and if you tell
4 them, you know, gee, having looked at it, there's a lot of good
5 stuff here, then there's some problem with the process.

6 I'm not saying 40,000 is the cutoff -- I can't really
7 determine that -- and I invite both sides' experts to tell me
8 if I've gotten this wrong but I've sat through a lot of
9 training sessions on this, wherever that cliff is, that where
10 is where the break should be. So if that was the only problem
11 you had with that part of the predictive coding process, then
12 it sounds like you all can go down this road, all of this,
13 without prejudice to additional search as may be necessary and
14 additional processes as may be necessary.

15 So is that the only problem, Ms. Wipper, or is there
16 anything more?

17 MS. WIPPER: No, there's a dispute about the scope of
18 relevancy. What happened --

19 THE COURT: I've ruled on that. That's what we spent
20 the morning doing.

21 MS. WIPPER: OK.

22 THE COURT: So whatever rulings I gave on that are
23 going to apply to the emails as well. So any positions they
24 were taking in the ESI protocol are now going to have to be
25 revised, based on what I have done this morning, and similarly

SOUTHERN DISTRICT REPORTERS, P.C.

(212) 805-0300

1214KDASC

CONFERENCE

1 on your side.

2 MS. WIPPER: OK, and also I'd like to respond to
3 defense counsel's description of their proposal. I'd like DOAR
4 to respond and give you an overview, if we may, on our proposal
5 on predictive coding.

6 THE COURT: All right, though I guess I'd like to know
7 where it differs.

8 MS. WIPPER: Well, it's actually a direct response to
9 their proposal.

10 THE COURT: OK.

11 MS. WIPPER: So who am I going to hear from?

12 MR. NEALE: Paul Neale, your Honor.

13 THE COURT: Mr. Neale?

14 MR. NEALE: I actually think you pointed to exactly
15 the issue. We have not taken issue with the use of predictive
16 coding or, frankly, with the confidence levels that they have
17 proposed except for the fact that it proposes a limit -- the
18 ultimate result of 40,000 documents before we have seen any of
19 the results coming out of the system.

20 THE COURT: I've already said -- and I want to make
21 sure that defense counsel realizes it -- I'm not buying your
22 40,000 as a pig in a poke. I understand the concept, but where
23 that line will be drawn -- whether it's 40,000, 50,000, 60,000,
24 20,000 -- is going to depend on what the statistics show for
25 the results.

SOUTHERN DISTRICT REPORTERS, P.C.

(212) 805-0300

1214KDASC

CONFERENCE

1 MR. ANDERS: I guess, your Honor, that's why I stood
2 up before, because I wanted to ask you something. I understand
3 that that cliff line may be at 80,000 documents. The reason
4 why we picked the 40,000 is what we're trying to do is also
5 incorporate the cost element. We picked 200,000 as what we
6 think --

7 THE COURT: Proportionality requires consideration of
8 results as well as costs. And if stopping at 40,000 is going
9 to leave a tremendous number of likely highly responsive
10 documents unproduced, it doesn't work. Plus, of course once
11 you have the predictive coding run, the cost after that is how
12 much you're doing an eyes-on review of. And once you've weeded
13 out the privilege documents -- and I assume you either have the
14 502(d) order or you will be providing one for me to sign off
15 on, because I think in a case of this size, if you're not
16 agreeing to one, you're committing malpractice -- how much
17 money you spend thereafter is a result of how much you want to
18 know what's in the documents or, putting it perhaps a different
19 way, CYA. If the first 60,000 are clearly showing that they're
20 highly relevant but you're running out of money after 40,000,
21 don't review the other 20,000. That's up to you.

22 MR. ANDERS: We've considered that, your Honor, and I
23 think the attorney-eyes-only type of agreement or designation
24 may be appropriate here, because one of the concerns we have
25 is, some of the plaintiffs are now working for competitors. To

SOUTHERN DISTRICT REPORTERS, P.C.

(212) 805-0300

1214KDASC

CONFERENCE

1 the extent that they're seeing --

2 THE COURT: This is not a case where I assume, other
3 than on anecdotal, that there is going to be much need for the
4 individual plaintiffs to look at the documents. I'm sure you
5 can all work that out.

6 Now, unfortunately it's 1:00 o'clock. I'm happy to
7 have you come back. I've got a 2:00 o'clock, and there may be
8 a 3:30 from people who forgot to show up this morning and were
9 told to try to get their act together and get here this
10 afternoon. You can come back this afternoon, you can come back
11 in a day or two. I think we have made some good progress, and
12 I know that you're coming from further away than usual, so I'd
13 like to make the most use of your time.

14 What's your pleasure? You want to come back at 3:30
15 in the afternoon and use the time from now to then? You can
16 use the jury room.

17 MR. ANDERS: Maybe, your Honor. The only reason why I
18 say that is, tomorrow I am leaving the country for a week for a
19 family vacation, so I'm out of pocket for a week; I'll have
20 some email but not a lot. So, again, I don't want to impose on
21 everybody else, but that's my scheduling issue, so I'm not sure
22 how much we'll get done within the next week.

23 THE COURT: That's why I'm suggesting you maximize --
24 I don't know what time your flight home is -- well, you're in
25 Morristown.

SOUTHERN DISTRICT REPORTERS, P.C.

(212) 805-0300

Exhibit

F

Natural Language vs. Boolean Query Evaluation: A Comparison of Retrieval Performance

Howard Turtle
West Publishing Company
Eagan, MN 55123
turtle@research.westlaw.com

Abstract

The results of experiments comparing the relative performance of natural language and Boolean query formulations are presented. The experiments show that on average a current generation natural language system provides better retrieval performance than expert searchers using a Boolean retrieval system when searching full-text legal materials. Methodological issues are reviewed and the effect of database size on query formulation strategy is discussed.

1 Introduction

Experimental results dating back to the 1970's have shown that simple natural language query techniques are at least as effective as Boolean techniques on comparable retrieval problems [1]. Despite these results, until recently all major commercial search services offered only Boolean query access to their collections. There are many reasons why commercial services were slow to adopt natural language queries, but one major reason was the size of the test collections on which the experimental results were based. Results based on a test collection of a few megabytes were not convincing when commercial search services routinely offered access to collections of many tens of gigabytes [2, 3]. Commercial services were aware that the difficulty of Boolean query formulation increased with collection size and were not convinced that natural language queries would scale up any better than Boolean.

Natural language searching has recently become available on WESTLAW, a large commercial system that provides access to U.S. legal materials [4, 5, 6, 7]. This system is a commercial implementation of the inference network retrieval model described in [8, 9]. Natural language queries now account for a significant portion of all WESTLAW searching (users can easily switch between Boolean and natural language) and, by implication, for a significant portion of all online access to U.S. legal materials.

For the purposes of this paper, a natural language search system is one that accepts as input a description of an information need in some natural language (such as English) and produces as output a list of documents ranked by the likelihood that they will be judged to match the information need. A Boolean search system accepts Boolean queries as input and uses the normal interpretation of first-order logic augmented with proximity operators to select a set of documents that satisfy the query. As discussed later, the set of documents retrieved in a Boolean search may be ranked. Retrieval models that accept Boolean queries but do not use first-order logic (e.g., [10, 11, 8]) are not considered. All commercial Boolean retrieval systems use first-order logic.

Development of the WESTLAW product involved direct comparison of natural language and Boolean query performance on test collections ranging in size from 250 megabytes to 6 gigabytes. In these tests, natural language results are significantly better than those obtained using "equivalent" Boolean queries formulated by expert searchers. These tests also show that natural language queries do, in fact, scale up better than Boolean – a natural language query offers similar performance across a wide range of collection sizes whereas a Boolean query must be refined when moved from a small collection to a large collection in order to produce result sets of manageable size.

What follows describes the methodology used to compare the performance of natural language and Boolean queries (Section 2) and presents and discusses the results of the experiments (Sections 3 and 4).

2 Methodology

In order to compare the retrieval performance of natural language and Boolean queries a set of natural language issue statements was developed by West attorneys. These issue statements represent the kinds of legal problems that lawyers would research using cases decided in the U.S. Federal Courts. These

Natural language:
What factors are important in determining what constitutes a vessel for purposes of determining liability of a vessel owner for injuries to a seaman under the Jones Act (46 USC 688)?
Boolean:
(741 +3 824) FACTOR ELEMENT STATUS FACT /P VESSEL SHIP BOAT /P (46 +3 688) "JONES ACT" /P INJUR! /S SEAMAN CREWMAN WORKER
Modified Boolean:
"JONES ACT VESSEL" (MEAN*** /3 VESSEL) (741 +3 824) (STATUS CONSTITUT! DETERMIN! DEFINITION /5 VESSEL) /P "JONES ACT" (46 +3 688) /P INJUR! /S SEAMAN CREW! WORKER WORKMAN

Figure 1: Sample queries

	FSupp	Fed
collection size (docs)	11,953	410,883
collection size (Mb)	254	6,053
unique stems	196,707	1,465,675
maximum stem frequency	11,947	336,000
stem occurrences	21,798,833	507,319,045
max within document frequency	1,309	2,119
mean document length (words)	1,823	1,235
queries	44	44
docs judged relevant per query	35	56
docs judged nonrel. per query	134	27

Table 1: Collection Characteristics

statements are generally a single sentence consisting of an average of just over 17 words. A representative query is shown in Figure 1.

A collection of cases decided in the Federal courts was drawn as a random sample from West's Federal Supplement database. This sample consists of roughly 250 megabytes of text and 12,000 federal cases. The characteristics of this FSupp collection are shown in Table 1.

A group of expert searchers was then asked to formulate Boolean queries for each of the natural language issue statements. The expert searchers were West editors and reference attorneys – all are attorneys and all have considerable experience in searching WESTLAW. Searchers were asked to produce the “best” query that they could for each issue statement. During query formulation they were allowed to conduct as many searches of the test collection as desired. Searchers were allowed to expand the query to include information that was not contained in the original issue statement. Searchers spent an average of 12.5 minutes formulating each Boolean query and conducted an average of 6.7 iterations to arrive at the final query form. A sample Boolean query is shown in Figure 1.

In many experimental studies, Boolean results are assumed to be unordered. In most commercial systems, Boolean results are ranked (ordered) using some document characteristic that is query independent (e.g., by date, alphabetically by author or title, or in some classification order). The Boolean results reported here are ranked in reverse chronological order, which is the presentation order used for virtually all Boolean searches on WESTLAW. Chronological or date ranking is surprisingly effective for legal materials where later decisions often interpret or supersede earlier decisions. Indeed, many users retrieve documents using relevance ranking, then sort the top *n* documents by date.

The result sets for these Boolean queries contained an average of 22.5 documents (the median result set contains 16 documents). All documents retrieved were evaluated for relevance by attorneys (not necessarily the attorney who formulated the query) on a three point scale: *on point*, *relevant*, and *not relevant*. *On point* cases address exactly the issue raised in the original issue statements, *relevant* cases

deal less directly with the issue, but are still considered to be relevant. For the purposes of evaluation, documents rated as *on point* or *relevant* were combined to form the set of relevant documents.

The natural language issue statements were used as input to a number of different commercial and prototype search engines. Many of these tests were conducted under the terms of non-disclosure agreements, but tests were also conducted using the INQUERY software at the University of Massachusetts, the SMART system at Virginia Polytechnic Institute, and using the Personal Librarian software developed by Personal Library Systems. For each natural language search, the top twenty documents were captured and rated for relevance (if they had not been previously rated). The union of all judgments from all experiments was used in the evaluation. On average, 169 documents have been rated for each query of which 35 were judged to be relevant.

The retrieval algorithms were tested extensively with the FSUPP test collection. This collection is large by traditional test collection standards, but it is small when compared to many of the collections available on WESTLAW. To test the effectiveness of the retrieval model on large collections a second test collection was selected from a database containing all U.S. Federal cases. This collection contains roughly 6 gigabytes of text and 410,000 documents. The documents in this collection have content similar to those in the FSUPP collection, but there are no documents common to the two collections. The characteristics of this Fed collection are shown in Table 1.

A new set of 91 queries was developed for this test collection, but the experiments reported here make use of the same query set developed for the FSUPP collection. Because the content of the cases in the two collections are similar, the original issue statements apply equally well to both collections.

2.1 Evaluation

The comparison of retrieval performance between ranked and unranked systems raises a number of difficult problems [1, 12, 13]. The two main problems are: 1) how can ranked versus unranked lists be compared, and 2) how can differences in the size of the returned set be reconciled?

With regard to the first problem, we take the view that the Boolean result is, in fact, ranked. The ranking is based on a document characteristic (date) that is independent of the query, but the ranking certainly is not random and users strongly prefer date ranking to an unordered presentation. While IR research has not yet produced experimental results that would allow us to quantify the degree to which users prefer good rankings to bad (i.e., we don't know how much an $x\%$ improvement in precision is worth to a typical user), the assumption that users prefer to see relevant documents first is plausible and pervasive, if unproven. Even in cases where the result is unordered we would argue that it is reasonable to compare rankings because the rankings determine what the user sees. Further, we are primarily interested in that portion of the ranking that an interactive user will see. For our purposes, a ranking that does well in the first twenty documents is to be preferred to a ranking that does poorly in the top twenty ranks but does very well at high recall levels.

A second problem in comparing the performance of Boolean and natural language queries arises because the techniques generally do not retrieve the same number of documents for a given query. The size of the sets returned by Boolean queries is highly variable. Some ranking algorithms also produce variable size results, but in WESTLAW, a fixed number of documents are retrieved for each natural language query. A user may request additional documents and may choose to alter the number of documents retrieved, but by default, twenty documents are returned. One solution to this problem is to use the size of the Boolean result set as the basis of comparison. If, for a given query, the Boolean result contains n documents then only the top n ranks for the natural language result are used for the comparison. There are several problems with this approach. First, it tends to favor the Boolean query and, in particular, it favors high precision Boolean queries. An adversary that knew how the evaluation was to be done could construct an unbeatable Boolean query by identifying a single document that was certain to be judged relevant and then constructing a query that retrieved only that document. Since Boolean queries tend to favor precision over recall [14] fixing on the Boolean result size would bias the evaluation.

A second problem with using the Boolean result size is that the set is often much larger than any interactive user is prepared to browse. On small collections it may be possible to produce Boolean results that are of manageable size, but on large collections this is often not possible. For example, the mean Boolean result set size on the FSUPP collection was 22.5 documents. As will be discussed later, the same queries run on the Fed collection retrieved an average of 206 documents. Even after revision to improve performance on Fed the Boolean queries still retrieved an average of 126 documents and some queries

retrieved several hundred documents. These result sets are far larger than most users are willing to browse.

Another solution to the result set size problem is to pick an arbitrary rank cutoff that is close to what an interactive user is willing to browse and to restrict evaluation to those ranks. This approach favors the natural language query if the Boolean result retrieves fewer documents than the cutoff. The Boolean searchers could be told to always retrieve result sets of at least a given size, but this would introduce further biases.

No single evaluation tool provides results that are truly unbiased, so we used four different measures in our tests: precision at standard recall points, raw precision and recall at a fixed rank cutoff, the relative performance of individual queries in the test set, and precision at all ranks up to some maximum value.

Average precision. We use an eleven point average precision with pessimistic interpolation [12] (see Table 3 for an example). For the Boolean queries we assume that the entire collection is ranked. The actual Boolean result fills in the top ranks for each query. All unretrieved documents are sorted by date and inserted in the ranks below the retrieved set, newest document first. Average precision is widely used and allows our results to be compared with earlier work but it tends to favor the natural language result at high recall levels because documents that are not retrieved by the Boolean query tend to appear well down in the ranking. Because the Boolean queries tend to exhibit relatively low recall, the precision values in the lower portion of the precision/recall table are probably not meaningful, especially for the FSupp collection where the Boolean results are small and the relevance judgments are fairly complete.

Raw precision and recall. A fixed rank cutoff of 20 documents is used. This number was originally selected because it is slightly smaller than the mean Boolean result on FSupp and because it is close to the number of documents that a typical user is willing to browse. We report raw precision (the percentage of documents in the top 20 that are relevant) and raw recall (the percentage of all relevant documents that are contained in the top 20). These numbers are easy to interpret since they are a direct indication of the number of relevant documents in the top 20, but they ignore the effect of ranking within the top 20.

We also report the maximum possible precision and recall figures that could be achieved given the set of relevance judgments. If every query had at least 20 relevant documents the maximum possible precision would be 100%. If every query had twenty or fewer relevant documents the maximum possible recall would be 100%.

Individual query performance. In addition to averages for the entire query set, we also report the number of queries for which one technique is superior. For each query we look at both the average precision and the top 20 precision. If both of these measures favor one technique by 10% or more, that technique is declared to be superior. If either difference is less than 10% or if each measure favors a different technique then the difference is declared to be insignificant.

Precision in the top n ranks. We will occasionally report raw precision for all ranks up to some fixed value. This is similar to top twenty precision, but it allows finer comparison. When averaged over a large number of queries precision values tend to monotonically decrease as rank increases. When the performance of two techniques is similar it is often possible to determine where in the ranking the precision curves cross.

2.2 Relevance judgments

It is generally not possible to assemble exhaustive relevance judgments for any but the smallest test collections. As a result, one major variable that affects nearly all information retrieval research is the quality and exhaustivity of the relevance judgments used. The most common technique for assembling relevance judgments is “pooling” in which actual results from several different retrieval tests are used to select documents to be judged for relevance [15, 16]. For the results reported here, all previously unevaluated documents in the top twenty ranks are pooled and submitted for evaluation. The only exception to this policy is that all documents retrieved by Boolean queries on FSupp have been evaluated regardless of the result size. The use of a twenty document cutoff is small when compared to the TREC evaluation where a cutoff of 100 is used [17]. The smaller cutoff reflects our focus on interactive retrieval rather than retrieval to support routing and filtering applications.

The FSupp collection has been used extensively at West Publishing and by a number of other research groups and the relevance judgments represent a pooling of the results of a large number of experiments using different retrieval techniques (including several relevance feedback techniques). While we have made no attempt to estimate the number of relevant but undetected documents that remain in the collection, the FSupp relevance judgments are “reasonably” complete. New experiments often turn up unevaluated

	NL better	No difference	Boolean better
FSupp	36	3	5
Fed	33	6	5
Fed - enhanced queries	29	8	7

Table 2: Per Query performance of natural language vs. Boolean queries

documents, but most of these are judged to be nonrelevant. The set of known relevant documents for the FSupp collection grows slowly.

The Fed collection is much newer and has not been used extensively outside of West. As a result, the number of documents that have been evaluated is much lower than for FSupp, both in terms of the number of documents per query (169 for FSupp versus 83 for Fed), and especially in terms of the percentage of the collection evaluated for each query. The number of relevant documents in the Fed collection is probably much larger than in FSupp and a reasonably high percentage of newly evaluated documents are judged to be relevant, so we regard the Fed judgments as relatively incomplete.

A second point with regard to relevance judgments is that experiments using standard test collections should report the number of unevaluated documents encountered in the top ranks. Most experiments report precision and recall figures that assume unevaluated documents are not relevant. If, however, a new ranking includes many unevaluated documents these results will generally be biased in favor of the retrieval technique(s) upon which the original relevance judgments are based. If any of the highly ranked but unevaluated documents are, in fact, relevant then the reported performance of the new ranking will be understated, sometimes substantially. A bit less obvious is the fact that the reported performance of the original ranking will generally be overstated if previously unevaluated documents are, in fact, relevant – the new relevant documents will occur relatively low in the original ranking which means that the original precision points occur at lower recall values. This shifts the precision/recall curve to the left which, in turn, reduces the area under the precision/recall curve.¹

The effect of incomplete relevance judgments can be quite pronounced. Early in the testing of West's natural language search engine we ran a series of experiments which showed a new technique produced a ranking that was 5% worse than the baseline in terms of average precision. The unevaluated documents in the top twenty ranks were pooled with other results and submitted to our reference attorneys for evaluation. With complete relevance judgments the new technique was determined to have produced a ranking that was, in fact, 15% better than the baseline. Since that time, one of the factors we examine in experimental results is the number of unevaluated documents in the top n ranks which gives a direct indication of how much the reported precision could be understated.

All of this suggests that test collections with fixed but incomplete sets of relevance judgments have a built-in bias that favors the techniques used to create the relevance judgments. The only real solution to this problem is to create mechanisms to allow documents to be evaluated on an ongoing basis. This is exactly the approach that has been adopted as part of the TREC evaluations [17] and is the approach used at West and in several other industrial research groups.

3 Results

3.1 FSupp Results

The baseline results on the FSupp collection are shown in Table 3. In terms of average precision, the natural language queries perform significantly better than the Boolean queries at all but the lowest recall level. When comparing precision at fixed rank cutoffs, Boolean queries retrieved a relevant document as the highest ranked document more often than did natural language (93.2% precision at rank one versus 84.1% for natural language). At ranks 2 and higher, however, the precision of the natural language query equals or exceeds that of the Boolean query. As discussed earlier, the low precision at high recall levels for Boolean queries arises in part because of the way that the Boolean rankings are constructed – if a document is not retrieved by the Boolean query it is likely to be assigned a relatively low rank.

Focusing on the top twenty documents retrieved, 42.3% of the documents retrieved by the Boolean queries are relevant whereas 57.0% of the documents retrieved by the natural language queries are relevant,

¹ Average precision at standard recall levels approximates the area under the precision/recall curve.

Recall	Precision (% change) - 44 queries		
	Boolean	Natural Language	
0	98.1	92.8	(-5.4)
10	77.9	87.4	(+12.2)
20	67.6	80.5	(+19.1)
30	39.7	71.9	(+81.3)
40	27.8	67.3	(+142)
50	14.9	62.7	(+321)
60	8.5	54.5	(+541)
70	6.5	47.3	(+627)
80	2.5	38.6	(+1471)
90	2.4	25.9	(+977)
100	0.3	13.7	(+4328)
average	31.5	58.4	(+85.6)
P20	42.3	57.0	(+34.9)
Max P20	87.3	87.3	
R20	24.4	32.9	(+34.9)
Max R20	50.3	50.3	

Table 3: Retrieval Performance on FSupp

an improvement of nearly 35% (given the set of relevance judgments, the maximum top twenty precision that could be achieved is 87.3%). These precision figures translate directly into top twenty recall figures of 24.4% for Boolean queries and 32.9% for natural language with a maximum possible recall of 50.3%.

Looking at individual queries (Table 2), natural language outperformed Boolean on 36 queries, Boolean outperformed natural language on 5 queries, and there was no significant difference on 3 queries.

Recall	Precision (% change) - 44 queries	
	Boolean	nl
0	85.1	94.0
10	76.1	84.7
20	67.7	80.4
30	54.7	72.3
40	47.5	65.0
50	37.3	54.5
60	33.1	43.1
70	26.8	35.3
80	21.1	27.1
90	13.2	18.8
100	6.2	7.4
average	42.6	53.0
P20	61.1	75.9
Max P20	97.8	97.8
R20	21.7	26.9
Max R20	34.7	34.7

Table 4: Retrieval performance on Fed

3.2 Fed results

When the same Boolean query sets are run on the larger Fed collection the number of documents retrieved increases substantially. The mean result set size for the FSupp collection is 22.5 documents whereas the mean results set size on the Fed collection for the same query set is 206 documents (median of 144 documents). Increasing the collection size by a factor² of roughly 24 increases the mean result set size by a factor of roughly 9.

²This ratio is based on the number of bytes of text. Fed contains roughly 34 times as many documents as FSupp.

The performance of these query sets on the Fed collection is shown in Table 4. Because of the large Boolean result set sizes and the relatively incomplete relevance judgments, the precision values are more nearly comparable at all standard recall levels than they are on FSupp. In effect, the queries that were formulated for the FSupp collection become high recall searches when moved to Fed. Most of the documents retrieved by the natural language queries are contained somewhere in the Boolean result set although the density of relevant documents in the top ranks is higher for natural language than for Boolean queries. In terms of average precision, natural language queries improve performance by 24.2%.

Using only the top twenty documents, the Boolean queries achieved 61.1% precision and 21.7% recall (maximum possible values are 97.8% precision and 34.7% recall). The natural language queries achieved 75.9% precision and 26.9% recall, an improvement of 24.2%.

Recall	Precision (% change) - 44 queries		
	Boolean	nl	
0	91.7	94.0	(+2.6)
10	80.3	84.7	(+5.5)
20	70.2	80.4	(+14.5)
30	62.2	72.3	(+16.3)
40	51.3	65.0	(+26.5)
50	38.2	54.5	(+42.5)
60	31.2	43.1	(+38.2)
70	20.6	35.3	(+70.8)
80	17.3	27.1	(+56.7)
90	11.2	18.8	(+67.4)
100	3.4	7.4	(+119.8)
average	43.4	53.0	(+22.0)
P20	69.2	75.9	(+9.7)
Max P20	97.8	97.8	
R20	24.5	26.9	(+9.7)
Max R20	34.7	34.7	

Table 5: Retrieval Performance on Fed with modified queries

With a mean result set size of over two hundred, it was clear that attorneys would not examine all of the cases retrieved by the Boolean queries. West's reference attorneys were asked to review the original queries and to revise any that they thought were inappropriate for use on the Fed collection. Seven of the original 44 queries were retained, the remaining 37 were modified. We had expected that the modifications would be to further restrict the Boolean queries developed on FSupp, but this was not always the case. Many of the queries were altered so that they would retrieve additional documents that would have been excluded by the initial formulation. Overall, the effect was to reduce the mean result set size from 206 documents to 126 documents. This result set size is still too large to allow most users to review all of the cases, but the reference attorneys felt that restricting the queries further would produce unacceptably low recall.

The modified Boolean queries do perform better than the originals, but still not as well as the natural language queries (Table 5). The modified Booleans have slightly higher average precision than the original Booleans. The improvement is significant at the low recall levels but performance falls off at high recall levels, primarily because modifications to reduce the result set size have also lowered recall. Looking at the top twenty documents, the modified Boolean queries have improved both precision and recall when compared to the original Booleans, but the natural language results still contain nearly 10% more relevant documents than the enhanced Boolean queries.

Looking at individual queries (Table 2), natural language outperformed the enhanced Booleans on 29 queries, Boolean outperformed natural language on 7 queries, and there was no significant difference on 8 queries.

4 Conclusion

Despite the methodological problems associated with comparing the performance of Boolean and natural language searching, the results of these experiments clearly support the assertion that, at least for full

text court opinions, natural language searching will present users with more relevant documents and will present these documents to the user in ranks that are likely to be browsed. On all of the evaluation metrics discussed here, natural language searches consistently produced better rankings. West has tested these techniques extensively on other material types (e.g., statutes, law review articles, administrative codes). Natural language searching performs consistently better than Boolean for all of the material types tested.

These performance comparisons are based on Boolean queries formulated by expert searchers that are intimately familiar with the collections involved, which suggests that the improvements over an “average” or novice searcher would be even more pronounced. Further, the Boolean queries were developed iteratively and incorporate information that is not contained in the natural language queries. The natural language queries were run without modification even though natural language queries can also be iteratively improved.

These results also point up some of the problems associated with searching large full-text collections using Boolean queries. As the size of the collection increases it becomes increasingly difficult to formulate a Boolean query that will produce a manageable result set. The natural language results are surprisingly consistent across the two collections. On both collections, natural language searches achieve 80 – 90% precision at the top ranks with average precision of over 50%. In some ways, natural language performance improves in large collections. Large collections often have more relevant documents for a given query so that precision in the top ranks actually increases with collection size. On FSUPP just over half (57%) of all documents in the top 20 ranks were judged to be relevant. On Fed, three out of every four documents in the top twenty ranks were judged to be relevant.

Despite the strong performance of natural language searching, Boolean query languages will not disappear anytime soon. For some queries and some kinds of materials Boolean techniques give better results. Further, some users prefer to use Boolean queries. Commercial systems will need to support both query types to be successful. At the same time, for the average searcher a natural language query will, on average, give better results than a Boolean query. Over time the performance differential can be expected to increase – there is a lot of room for improvement of current generation natural language systems.

References

- [1] Karen Sparck Jones. Retrieval system tests 1958-1978. In Karen Sparck Jones, editor, *Information Retrieval Experiment*, chapter 12, pages 213–255, Butterworth, 1981.
- [2] Donald R. Swanson. Information retrieval as a trial and error process. *Library Quarterly*, 47(2):128–148, 1977.
- [3] David C. Blair. *Language and Representation in Information Retrieval*. Elsevier Science Publishers, Amsterdam, 1990.
- [4] Laura Duncan. New research net drops computerese. *Chicago Daily Law Bulletin*, 2, September 16 1992.
- [5] IDP Report. West unveils new natural language interface. *Information and Database Publishing Report*, 13(35):1,6, September 18 1992.
- [6] Cary Griffith. WESTLAW’s WIN: not only natural, but new. *Information Today*, 9–11, October 1992.
- [7] Cary Griffith. WESTLAW’s winning ways. *Law Office Computing*, 31–38, February/March 1993.
- [8] Howard Turtle. *Inference Networks for Document Retrieval*. PhD thesis, Computer Science Department, University of Massachusetts, Amherst, MA 01003, 1990. Available as COINS Technical Report 90-92.
- [9] Howard Turtle and W. Bruce Croft. Evaluation of an inference network-based retrieval model. *ACM Transactions on Information Systems*, 9(3):187–222, July 1991.
- [10] Edward A. Fox. *Extending the Boolean and Vector Space Models of Information Retrieval with P-Norm Queries and Multiple Concept Types*. PhD thesis, Cornell University, 1983.

- [11] Edward A. Fox and Mathew B. Koll. Practical enhanced Boolean retrieval: experiences with the SMART and SIRE systems. *Information Processing and Management*, 24(3):257–267, 1988.
- [12] E. M. Keen. Evaluation parameters. In Gerard Salton, editor, *The SMART Retrieval System – Experiments in Automatic Document Processing*, chapter 5, pages 74–111, Prentice Hall, 1971.
- [13] Lynn Evans. An experiment: search strategy variation in SDI profiles. In Karen Sparck Jones, editor, *Information Retrieval Experiment*, chapter 14, pages 285–315, Butterworth, 1981.
- [14] D. C. Blair and M. E. Maron. An evaluation of retrieval effectiveness for a full-text document retrieval system. *Communications of the ACM*, 28(3):290–299, 1985.
- [15] Karen Sparck Jones and C. J. van Rijsbergen. *Report on the Need for and the Provision of an ‘Ideal’ Information Retrieval Test Collection*. Technical Report British Library Research and Development Report 5266, Computer Laboratory, University of Cambridge, 1975.
- [16] Karen Sparck Jones and R. G. Bates. *Report on a Design Study for the ‘Ideal’ Information Retrieval Test Collection*. Technical Report British Library Research and Development Report 5428, Computer Laboratory, University of Cambridge, 1977.
- [17] Donna Harman. Overview of the first text retrieval conference (TREC-1). In Donna K. Harman, editor, *The First Text Retrieval Conference (TREC-1)*, pages 1–20, National Institute of Standards and Technology, March 1993. Proceedings available as NIST Special Publication 500-207.

Exhibit

G



SEARCH, FORWARD

Will manual document review and keyword searches be replaced by computer-assisted coding?

First, there was manual review — the “traditional” method of document review. As a young associate at a major New York law firm in the late 1970s, I reviewed boxes of files for relevance, “hot documents,” and privilege. To gather the paper documents, you went to the client and asked where they kept files about “X” (“X” being the issue(s) involved in the lawsuit). Often there was a central file labeled “X,” and employees kept their own working files as well. Occasionally, you had to go to the dreaded warehouse, where boxes might not be indexed, and working conditions always were less than ideal.

Review was linear. There was no way to deduplicate documents or organize them by types. You reviewed whatever box landed on your desk; colleagues might be reviewing a carbon copy of the same file. Hopefully, you both coded it the same. (Even today, it is not unusual for a document to be produced while another copy is on the privilege log.)

When associate billing rates became too high, firms turned to paralegals, staff attorneys, or contract attorneys. Whether this had any effect on the quality of the review was beside the point; economics drove the change.

Despite its flaws, many senior lawyers (and some clients) still consider manual review to be the “gold standard” against which other review techniques are compared. While the volume of electronically stored information (and concomitant expense) has largely eliminated manual review as the sole method of document review, manual review remains used along with, for example, keyword screening. Let us consider whether manual review as the gold standard is myth or reality.

Two recent research studies clearly demonstrate that computerized searches are at least as accurate, if not more so, than manual review. Herb Roitblatt, Ann Kershaw, and Patrick Oot, of the Electronic Discovery Institute, concluded that “[o]n every measure, the performance of the two computer systems was at least as accurate (measured against the original review) as that of human re-review.” (“Document Categorization in

Legal Electronic Discovery: Computer Classification vs. Manual Review,” *Journal of Am. Society for Information Science & Technology*, 61(1):70-80 (2010).)

Likewise, Wachtell, Lipton, Rosen & Katz litigation counsel Maura Grossman and University of Waterloo professor Gordon Cormack, using data from the Text Retrieval Conference Legal Track, concluded that “[T]he idea that exhaustive manual review is the most effective — and therefore the most defensible — approach to document review is strongly refuted. Technology-assisted review can (and does) yield more accurate results than exhaustive manual review, with much lower effort. (“Technology-Assisted Review in E-Discovery Can Be More Effective and More Efficient Than Exhaustive Manual Review,” *Richmond J. of Law & Tech.*, Vol. XVII, Issue 3, 1-48 (2011)).

Grossman and Cormack note that “not all technology-assisted reviews . . . are created equal” and that future studies will be needed to “address which technology-assisted review process(es) will improve most on manual review.”

KEY WORDS

Because the volume of ESI has made full manual review virtually impossible, lawyers have turned to keywords to cull ESI (particularly e-mail) for further (manual) review. A basic problem is that absent cooperation, the way most lawyers engage in keyword searches is, as Ralph Losey suggests, the equivalent of “Go Fish.” The requesting party guesses which keywords might produce evidence to support its case without having much, if any, knowledge of the responding party’s “cards” (i.e., the terminology used by the responding party’s custodians). Indeed, the responding party’s counsel often does not know what is in its own client’s “cards.”

The problems with keyword search are well known. Lawyers are used to doing keyword searches in “clean” databases, such as Westlaw and Lexis, which use full sentences, full words (not abbreviations), and largely the same words to describe the same concept. E-mail collections are not clean databases. People use different words to describe the same concept; even business

e-mails are informal, rampant with misspellings, abbreviations, and acronyms.

The object of search is to produce high recall and high precision. Recall is the fraction of relevant documents identified during a review, i.e., a measure of completeness. Precision is the fraction of identified documents that are relevant, i.e., it is a measure of accuracy or correctness.

When keywords return false positives — documents that have the keywords but are not relevant — the responding party has to use expensive manual review to find the truly relevant documents. It is not uncommon for a poorly chosen keyword to return more “junk” than responsive documents, i.e., low precision. The goal of search is to produce high recall and high precision (in a cost-effective way).

How effective is keyword searching? In 1985, scholars David Blair and M.E. Maron collected 40,000 documents from a Bay Area Rapid Transit accident, and instructed experienced attorney and paralegal searchers to use keywords and other review techniques to retrieve at least 75% of the documents relevant to 51 document requests. Searchers believed they met the goals, but their average recall was just 20%. This result has been replicated in the TREC Legal Track studies over the past few years.

Judicial decisions have critiqued keyword searches. Important early decisions in this area came from magistrate judges John Facciola (District of Columbia) and Paul Grimm (Maryland). See *United States v. O’Keefe*, 37 F. Supp. 2d 14, 24 (D.D.C. 2008) (Facciola, M.J.); *Equity Analytics, LLC v. Lundin*, 248 F.R.D. 331, 333 (D.D.C. 2008) (Facciola, M.J.); and *Victor Stanley, Inc. v. Creative Pipe, Inc.*, 250 F.R.D. 251, 260, 262 (D. Md. 2008) (Grimm, M.J.).

I followed their lead with *William A. Gross Construction Associates, Inc. v. American Manufacturers Mutual Insurance Co.*, 256 F.R.D. 134, 136 (S.D.N.Y. 2009) (Peck, M.J.).

“This Opinion should serve as a wake-up call to the Bar in this District about the need for careful thought, quality control, testing, and cooperation with opposing counsel in designing search

terms or ‘keywords’ to be used to produce e-mails or other electronically stored information (‘ESI’),” I wrote.

My opinion concluded: “Electronic discovery requires cooperation between opposing counsel and transparency in all aspects of preservation and production of ESI. Moreover, where counsel are using keyword searches for retrieval of ESI, they at a minimum must carefully craft the appropriate keywords, with input from the ESI’s custodians as to the words and abbreviations they use, and the proposed methodology must be quality control tested to assure accuracy in retrieval and elimination of ‘false positives.’ It is time that the Bar — even those lawyers who did not come of age in the computer era — understand this.”

Despite these (and other) judicial criticisms of the use of keywords without sufficient testing and quality control, many counsel still use the “Go Fish” model of keyword search. Cooperation is important, but without testing and quality control cooperation alone is not the answer.

COMPUTER-ASSISTED SEARCH

Even with keyword searching, lawyers have turned to certain computer-assisted approaches to further reduce review cost. Boolean connectors can be used (such as “and,” “or,” “w/in,” “but not”). In addition, deduplicating the ESI (either within a single custodian or across the entire production) greatly reduces both volume and the chance of the same e-mail being coded differently by different reviewers. Grouping “near duplicates” takes that a step further. Threading e-mail chains is another useful technique.

If the hot topic in 2010 conferences was proportionality, this year it is computer-assisted coding, often generically called “predictive coding.” By computer-assisted coding, I mean tools (different vendors use different names) that use sophisticated algorithms to enable the computer to determine relevance, based on interaction with (i.e., training by) a human reviewer.

Unlike manual review, where the review is done by the most junior staff, computer-assisted coding involves a senior partner (or team) who review and code a “seed set” of documents. The computer identifies properties of those documents that it uses to code other documents. As the senior reviewer continues to code more sample documents, the computer predicts the reviewer’s coding. (Or, the computer codes some documents and asks the senior reviewer for feedback.)

When the system’s predictions and the reviewer’s coding sufficiently coincide, the system has learned enough to make confident predictions for the remaining documents. Typically, the senior lawyer (or team) needs to review only a few thousand documents to train the computer.

Some systems produce a simple yes/no as to relevance, while others give a relevance score (say, on a 0 to 100 basis) that counsel can use to prioritize review. For example, a score above 50 may produce 97% of the relevant documents, but constitutes only 20% of the entire document set.

Counsel may decide, after sampling and quality control tests, that documents with a score of below 15 are so highly likely to be irrelevant that no further human review is necessary. Counsel can also decide the cost-benefit of manual review of the documents with scores of 15-50.

To my knowledge, no reported case (federal or state) has ruled on the use of computer-assisted coding. While anecdotally it appears that some lawyers are using predictive coding technology, it also appears that many lawyers (and their clients) are waiting for a judicial decision approving of computer-assisted review.

Perhaps they are looking for an opinion concluding that: “It is the opinion of this court that the use of predictive coding is a proper and acceptable means of conducting searches under the Federal Rules of Civil Procedure, and furthermore that the software provided for this purpose by [insert name of your favorite vendor] is the software of choice in this court.” If so, it will be a long wait.

Judicial decisions, including *Victor Stanley*, *O’Keefe* and *Gross*, are highly critical of the keywords used by the parties. These decisions did not “endorse” or “approve” of keyword searching. Nevertheless, lawyers seem to believe that the judiciary has signed off on keywords, but has not on computer-assisted coding.

In addition to reluctance to be the guinea pig for a decision on predictive coding, lawyers perhaps are concerned that they will have to go through a *Daubert* hearing as to the “admissibility” of the results of predictive coding. Perhaps this fear comes from *O’Keefe*, where Judge Facciola said that opining on what keyword is better “is truly to go where angels fear to tread,” and is a topic “beyond the ken of a layman and requires that any such conclusion be based on evidence that, for example, meets the criteria of Rule 702 of the Federal Rules of Evidence,” dealing with expert opinions.

Lawyers’ fears in this regard seem largely misplaced. First, Facciola’s comments were directed at keywords, but everyone is using keywords, and I know of no decision after *O’Keefe* requiring expert testimony as to the use of keywords.

Second, with due respect to Facciola, I do not think *Daubert* applies — it applies when

an expert will testify at trial in order to admit into evidence opinions or results (e.g., the result of DNA testing reveals a match).

Here, the hundreds of thousands of e-mails produced are not being offered into evidence at trial as the result of a scientific process. Rather, whether the handful of e-mails offered as trial exhibits is admissible is dependent on the document itself (e.g., whether it is a party admission or a business record), not how it was found during discovery.

That said, if the use of predictive coding is challenged in a case before me, I will want to know what was done and why that produced defensible results. I may be less interested in the science behind the “black box” of the vendor’s software than in whether it produced responsive documents with reasonably high recall and high precision.

That may mean allowing the requesting party to see the documents that were used to train the computer-assisted coding system. (Counsel would not be required to explain why they coded documents as responsive or non-responsive, just what the coding was.) Proof of a valid “process,” including quality control testing, also will be important.

Additionally, counsel can point to the TREC study and other reported studies that generally show that computer-assisted coding technology works at least as well if not better than keywords or manual review.

Of course, the best approach to the use of computer-assisted coding is to follow the Sedona Cooperation Proclamation model. Advise opposing counsel that you plan to use computer-assisted coding and seek agreement; if you cannot, consider whether to abandon predictive coding for that case or go to the court for advance approval.

Until there is a judicial opinion approving (or even critiquing) the use of predictive coding, counsel will just have to rely on this article as a sign of judicial approval. In my opinion, computer-assisted coding should be used in those cases where it will help “secure the just, speedy, and inexpensive” (Fed. R. Civ. P. 1) determination of cases in our e-discovery world.

Andrew Peck is a United States magistrate judge for the Southern District of New York. E-mail: Andrew.Peck@nysd.uscourts.gov.

Reprinted with permission from the October 2011 edition of LAW TECHNOLOGY NEWS. © 2011 ALM Media Properties, LLC. All rights reserved. Further duplication without permission is prohibited. For information, contact 877-257-3382 or reprints@alm.com. #010-10-11-01